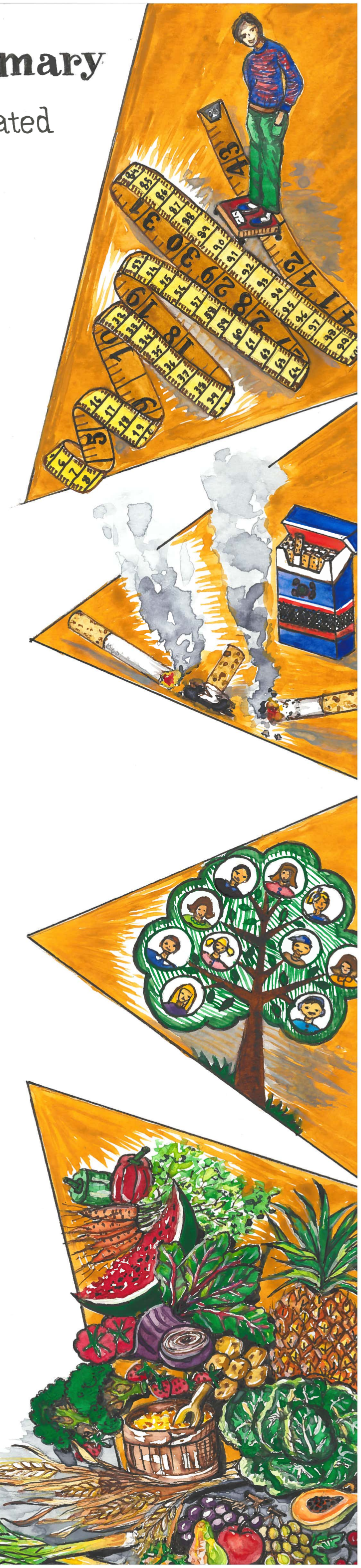
Cancer of Unknown Primary

uncovering risk factors associated with a neglected disease



Karlijn Hermans

Cancer of Unknown Primary

uncovering risk factors associated with a neglected disease

Karlijn Hermans

Cancer of Unknown Primary
uncovering risk factors associated with a neglected disease

ISBN: 978-94-6458-504-9
Cover design: Maria da Graça de Jesus da Cruz Monteiro
Layout: Karlijn Hermans
Printed by: Ridderprint | www.ridderprint.nl

Cancer of Unknown Primary
uncovering risk factors associated with a neglected disease

DISSERTATION

To obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus, Prof. dr. P. Habibović,
in accordance with the decision of the Board of Deans,
to be defended in public on Tuesday, 15th of November 2022, at 13:00 hours

By

Karlijn Elisabeth Petrus Esther Hermans
Born 6th of May 1994, Roermond, the Netherlands

Promotor

Prof. dr. ir. P.A. van den Brandt

Co-promotores

Dr. L.J. Schouten

Dr. ing. C. Loef (Integraal Kankercentrum Nederland)

Assessment Committee

Prof. dr. V.C.G. Tjan-Heijnen *(Chair)*

Prof. dr. B. Kremer

Prof. dr. J.W.M. Muris

Prof. dr. V.E.P.P. Lemmens (Integraal Kankercentrum Nederland)

Prof. dr. V.T.H.B.M. Smit (Leiden University Medical Center)

The research presented in this thesis was conducted at GROW-School for Oncology and Developmental Biology, Department of Epidemiology, Maastricht University. This work was financially supported by Wereld Kanker Onderzoek Fonds Nederland (WCRF-NL), as part of the World Cancer Research Fund International grant program (grant number 2017/1628). Printing and dissemination of this thesis was financially supported by the Department of Epidemiology, Integraal Kankercentrum Nederland, and Missie Tumor Onbekend.

Contents

CHAPTER 1

GENERAL INTRODUCTION

Cancer of Unknown Primary (CUP) is a metastatic cancer with no identifiable primary tumour origin (1-4). In most cancer cases, there is a clear onset of the primary tumour and its progression, but sometimes metastases are the first symptom while the primary tumour cannot be found despite the completion of initial diagnostic workup and histological and/or cytological verification. In general, a cancer diagnosis has a major impact on the patients' life as well as that of their relatives. The uncertainty of cancer patients is even more complex for CUP patients, as it remains unclear what type of cancer it is and whether treatment(s) is possible.

Incidence and survival

CUP is a frequently diagnosed cancer and ranks fourth place in the most common metastatic forms of cancer (5, 6). Over the last 10-20 years the global CUP incidence has decreased from 3-5% to 1-2% of all cancers (7). This decline derives from multiple factors such as increased availability and improved diagnostic accuracy of advanced diagnostic imaging (7, 8); increased variety and accuracy of molecular tracers to detect the primary tumour origin; early detection of the primary tumour origin and metastasis by cancer screening which may support tumour regression; changes and agreement with respect to coding-criteria for recording cases in cancer registries (9); and possibly due to a lower prevalence of exposure to risk factors like smoking and alcohol consumption (2, 10).

Hitherto, survival patterns for CUP patients remain poor. The median survival ranges between three to ten months after the first hospital visit. Survival times are dependent on the histology of the malignancy, as well as the general condition of the patient (e.g. age and/or comorbidities) (11). Adequate timing with respect to diagnostic and pathology examination(s) is important as delays can postpone the start of the treatment(s) and, hence, affect its effectiveness (12).

CUP-definition

The United Kingdom's National Institute for Health and Clinical Excellence (NICE) guideline (2010) on cancer has categorised CUP according to three definitions 1) malignancy of undefined primary origin (MUO), 2) provisional CUP: metastatic epithelial or neuroendocrine malignancy identified based on microscopical verification (histology and/or cytology), and 3) confirmed CUP: metastatic epithelial or neuroendocrine malignancy identified based on final histology, with no primary

site detected despite a selected initial screen of investigations, specialist review, and specialised investigations as appropriate (9, 13). These NICE-criteria are particularly useful in clinical settings; however, population-based research datasets contain a mixture of CUP cases without a clear distinction between provisional or confirmed cases. The International Classification of Diseases for Oncology has defined CUP as an 'unknown primary site' if a tumour is metastatic and the primary site is unknown. This definition is more applicable for population-based research. Yet, this definition has been updated over time, which complicates the measurement of time trends. In addition, there are international differences for coding criteria that are applied by cancer registries. For example, in the Netherlands, the cancer clinical practice guidelines define CUP as metastasis of an unknown primary tumour origin, based on a cytological and/or histological proven metastasis of cancer (14). The mixture of CUP-definitions makes it very difficult to compare this entity between countries (7, 15). Overall, cancer registries pursue to register all cancer cases, for which data derive from pathology excerpts and hospital discharge papers. Accordingly, the quality and completeness of these data influence the possibility to correctly record cancer cases. Incomplete data could, therefore, result in a misdiagnosed CUP.

Advancements

As briefly introduced, decreased CUP occurrence may be a result of improved diagnostic imaging methods and molecular profiling (7, 12). In the Netherlands, positron emission tomography (PET) was introduced and implemented in 2000-2005 for detecting malignancies more accurately. Alongside the introduction of PET, there have been advances in specific serological tumour markers (16), immunohistochemical techniques (17-19), gene expression-based profiling (20), endoscopy (21), and magnetic resonance imaging (MRI) (22), computed tomography (CT) (16), and ^{18}F-FDG PET-CT (23, 24). The introduction of PET and the increased use of tumour marker analysis paralleled the decrease in CUP incidence, it may be that both techniques individually/combined contributed to the identification of primary tumour origins as well as metastases (2). In particular, the ^{18}F-FDG PET-CT method has proven to be valuable in identifying primary tumour origins (23, 24). In the Netherlands, another important asset is the use of a national population screening programme for cancers of the breast, cervix, and colorectum. This programme is intended for early detection and localisation of diseases, which enables more effective treatment(s) (25). Early detection may also influence CUP occurrence, as prompt treatment involves treating a primary

tumour before it can metastasise. Furthermore, whole-genome sequencing is found to be effective for identifying primary tumour origins based on genotyping in some CUP patients. The identification of these primary tumour origins may be utilised to stratify patients towards specific therapies for precision medicine in cancer (26, 27). In addition, mutational profiling is found to be useful for identifying specific molecular pathways which derive from patterns of exposures that are linked to cancer cellular mechanisms that may be used for therapies (28). Another important asset to improve the identification of primary tumour origins may lie in the assistive use of artificial intelligence (29). The advancements in diagnostic imaging and insights of the abovementioned novel techniques have great potential to improve the detection and prediction of tumour subtypes and subsequently guide treatment decisions for CUP patients (26, 28, 29). The disease occurrence may, therefore, continue to decrease in prospective years (2).

Unidentifiability of the primary tumour

Pathologists have described several mechanisms that may explain the inability of identifying the primary tumour origin, despite the completion of extensive diagnostic and/or pathology investigations; 1) the primary tumour may not be recognized as the primary tumour due to atypical histologic features, 2) there has been a malignant transformation of ectopic tissue; which was frequently seen for extragonadal germ cell tumours (if the germ cell phenotype of a malignant tumour was unclear, it may have been classified as CUP, 3) the primary tumour was removed before it became evident after having seeded metastases (history of genitourinary tract surgery or related interventions), 4) the primary tumour disappeared as a result of growth inhibition of the primary lesion due to response of the immune system, therefore, it may have shrunk until extinction, or 5) the primary lesion spread to one or several metastatic sites before becoming malignant (30). Nonetheless, it is also possible that other factors played a role in the decision to refrain from further diagnostic investigation such as age, comorbidities, performance status, localisation of metastasis, or the patient's decision. In general, cancer treatment(s) is targeted at the primary tumour origin. The inability of identifying the primary tumour, thus, complicates organ and tumour specific treatment possibilities (6).

Cancer risk factors and cancer prevention

Various risk factors are linked to increase a person's risk of cancer development. Overall, cancer prevention recommendations advise to not use any form of tobacco, maintain healthy body weight, be physically active, consume a healthy diet (including whole grains, pulses, vegetables and fruits, while limiting the intake of foods high in sugar or fat, sugary drinks, processed and red meats, and foods high in salt), and avoid alcohol consumption (31-38). A study has indicated that for cancer overall, approximately 42-50% of the cases could be prevented if exposures are controlled effectively (39). In 2018, the WCRF/AICR updated ten cancer prevention recommendations concerning lifestyle factors such as diet, nutrition and physical activity (31). Previous studies have demonstrated that adhering to these recommendations, is associated with lower cancer risk, and with lower total and cancer mortality (40, 41). However, whether adherence to these recommendations is also associated with CUP risk, has not been studied before.

Rationale and aim of this thesis

Despite the frequent occurrence and bleak prognosis of CUP, research into its risk factors remains particularly scarce. Consequently, it remains unclear whether CUP has a specific risk factor profile. The identification of risk factors is important to guide cancer prevention. Therefore, the aim of this thesis was to investigate the association between individual lifestyle components: 1) alcohol consumption, cigarette smoking, anthropometry, physical activity, vegetable and fruit consumption, meat consumption, family history of cancer, and diabetes mellitus in relation to CUP risk, and lifestyle as an overall component by studying 2) whether adherence to the WCRF/AICR lifestyle recommendations for cancer prevention is associated with CUP risk. Finally, we decided to combine our findings with the existing epidemiological evidence in an up-to-date comprehensive review.

Study design and population

Most studies were performed within the context of the Netherlands Cohort Study on diet and cancer (NLCS). This prospective cohort includes a study population of 120,852 participants (58,279 men and 62,573 women) aged 55-69 years at baseline in 1986. Participants originated from 204 Dutch municipal population registries (42). All participants completed a mailed, self-administered questionnaire on dietary habits and other cancer risk factors at baseline in 1986. The questionnaire was evaluated for

its validity and reproducibility (43, 44). Efficient data processing and analysis were achievable by applying a case-cohort approach. Subsequently, incident cancer cases were derived from the full cohort, while the number of person-years at risk for the full cohort was estimated from a subcohort of 5,000 participants who were randomly sampled from the full cohort at baseline in 1986 (42). The institutional review boards of the Netherlands Organization for Applied Scientific Research TNO (Zeist) and Maastricht University (Maastricht) approved the execution of the NLCS.

To generate a CUP-dataset, we identified incident CUP cases through annual record linkage of the full cohort with the Netherlands Cancer Registry (NCR) and the Dutch Pathology Registry (PALGA) (45). Participants were followed up for 20.3 years (from 17 September 1986 until 31 December 2006). Information regarding the site of the metastasis was obtained from the NCR, supplementary information was retrieved from PALGA pathology excerpts. CUP cases were categorised according to histology (adenocarcinoma, undifferentiated carcinoma, squamous cell carcinoma, neuro-endocrine carcinoma, and other carcinoma); according to the number of metastases (multiple metastases of the same type were counted as one metastatic site, e.g., bone metastases in hip and vertebra were counted as one); according to localisation of the metastasis (up to four locations), and according to survival duration (≤1 and >1 year after diagnosis). In our studies, CUP is defined as a metastasised epithelial malignancy with no identifiable primary tumour origin after cytological and/or histological verification during a patient's lifetime. This definition solely includes epithelial malignancies according to the International Classification of Diseases for Oncology version 3: M-8000 - M-8570. For this purpose, cases with sarcoma, lymphoma, mesothelioma, and melanoma were excluded.

Outline of the thesis

The following chapters of this thesis describe the investigation of the risk factors that were studied in association with CUP risk. First, individual lifestyle components are investigated: alcohol consumption and cigarette smoking in Chapter 2, anthropometry and physical activity in Chapter 3, vegetable and fruit consumption in Chapter 4, meat consumption in Chapter 5, family history of cancer in Chapter 6, and diabetes mellitus in Chapter 7. Second, lifestyle as an overall component by measuring adherence to the WCRF/AICR lifestyle recommendations for cancer prevention and CUP risk is described in Chapter 8. A comprehensive review on epidemiological findings of CUP risk factors is presented in Chapter 9. To conclude this thesis, the results and future recommendations are discussed in Chapter 10.

References

1. Loffler H, Kramer A. Introduction. In: Loffler H, Kramer A, editors. Cancer of Unknown Primary Springer International Publishing Switzerland; 2016.
2. Schroten-Loef C, Verhoeven R, de Hingh I, van de Wouw A, van Laarhoven H, Lemmens V. Unknown primary carcinoma in the Netherlands: decrease in incidence and survival times remain poor between 2000 and 2012. Eur J Cancer. 2018;101:77-86.
3. Hillen H. Unknown primary tumours. Postgraduate medical journal. 2000;76(901):690-3.
4. Hemminki K. Epidemiology, Risk Factors, and Survival in CUP: Pointers to Disease Mechanisms. In: Kramer A, Loffler, H, editor. Cancer of Unknown Primary. Switzerland: Springer International Publishing 2016.
5. Comprehensive Cancer Centre the Netherlands. Incidentie Primaire Tumor Onbekend: Integraal kankercentrum Nederland; 2020 [Available from: https://iknl.nl/kankersoorten/primaire-tumor-onbekend/registratie/incidentie].
6. Meijer L, De Peuter R, Van der Zwan JM, Loef C. Primaire Tumor Onbekend: Integraal kankercentrum Nederland; 2020 [Available from: https://www.iknl.nl/getmedia/ea952e4d-9af7-4d59-9f03-865c858bc0cd/Rapport_PrimaireTumorOnbekend_IKNL_2020.pdf].
7. Rassy E, Pavlidis N. The currently declining incidence of cancer of unknown primary. J Cancer Epidemiol. 2019;61:139-41.
8. Jerusalem G, Rorive A, Ancion G, Hustinx R, Fillet G. Diagnostic and therapeutic management of carcinoma of unknown primary: radio-imaging investigations. J Annals of Oncology. 2006;17:x168-76.
9. Brewster DH, Lang J, Bhatti LA, Thomson CS, Oien KA. Descriptive epidemiology of cancer of unknown primary site in Scotland, 1961-2010. Cancer Epidemiol. 2014;38(3):227-34.
10. Kaaks R, Sookthai D, Hemminki K, Kramer A, Boeing H, Wirfalt E, et al. Risk factors for cancers of unknown primary site: Results from the prospective EPIC cohort. Int J Cancer. 2014;135(10):2475-81.
11. Hemminki K. Epidemiology, Risk Factors, and Survival in CUP: Pointers to Disease Mechanisms. Cancer of Unknown Primary: Springer; 2016. p. 5-25.
12. Meijer L, Verhoeven RHA, de Hingh I, van de Wouw AJ, van Laarhoven HWM, Lemmens V, et al. Extensive diagnostic work-up for patients with carcinoma of unknown primary. Clinical & experimental metastasis. 2021;38(2):231-8.
13. National Institute for Health and Care Excellence. Metastastic malignant disease of unknown primary origin. Diagnosis and management of metastatic disease of unknown primary origin. NICE clinical guideline 104 2010. Available from: guidance.nice.org.uk/cg104
14. Oncoline. Primaire tumor onbekend 2012 [updated June 24th, 2012. Available from: https://www.oncoline.nl/primaire-tumor-onbekend.
15. Kok S, van Laarhoven H, Lemmens V, Schroten C. International variation in diagnostic and treatment guidelines for carcinoma unknown primary. Cancer. 2014;1:2015.
16. Iizuka Y, Iizuka H, Tsutsumi S, Nakagawa Y, Nakajima T, Sorimachi Y, et al. Diagnosis of a previously unidentified primary site in patients with spinal metastasis: diagnostic usefulness of laboratory analysis, CT scanning and CT-guided biopsy. European Spine Journal. 2009;18(10):1431-5.
17. Park JM, Jung CK, Choi YJ, Lee KY, Kang JH, Kim MS, et al. The use of an immunohistochemical diagnostic panel to determine the primary site of cervical lymph node metastases of occult squamous cell carcinoma. Human Pathology. 2010;41(3):431-7.

18. Strojan Flezar M, Srebotnik Kirbis I. Identification of carcinoma origin by thyroid transcription factor-1 immunostaining of fine needle aspirates of metastases. Cytopathology : official journal of the British Society for Clinical Cytology. 2009;20(3):176-82.
19. Morawietz L, Floore A, Stork-Sloots L, Folprecht G, Buettner R, Rieger A, et al. Vergleich immunhistologischer mit genexpressionsbasierter Diagnostik. Der Pathologe. 2009;30(2):168.
20. Monzon FA, Koen TJ. Diagnosis of metastatic neoplasms: molecular approaches for identification of tissue of origin. Archives of pathology & laboratory medicine. 2010;134(2):216-24.
21. Sakai A, Okami K, Ebisumoto K, Sugimoto R, Maki D, Iida M. New techniques to detect unknown primaries in cervical lymph node metastasis. The Laryngoscope. 2010;120(9):1779-83.
22. Buchanan CL, Morris EA, Dorn PL, Borgen PI, Van Zee KJ. Utility of breast magnetic resonance imaging in patients with occult primary breast cancer. Annals of surgical oncology. 2005;12(12):1045-53.
23. Dong M-j, Zhao K, Lin X-t, Zhao J, Ruan L-x, Liu Z-f. Role of fluorodeoxyglucose-PET versus fluorodeoxyglucose-PET/computed tomography in detection of unknown primary tumor: a meta-analysis of the literature. 2008;29(9):791-802.
24. Kwee TC, Kwee RM. Combined FDG-PET/CT for the detection of unknown primary tumors: systematic review and meta-analysis. European Radiology. 2009;19(3):731-44.
25. National Institute for Public Health and the Environment MoH, Welfare and Sport. Population screening programmes 2022 [Available from: https://www.rivm.nl/en/population-screening-programmes.
26. Priestley P, Baber J, Lolkema MP, Steeghs N, de Bruijn E, Shale C, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. Nature. 2019;575(7781):210-6.
27. Rassy E, Assi T, Pavlidis N. Exploring the biological hallmarks of cancer of unknown primary: where do we stand today? Br J Cancer. 2020:1-9.
28. Pentheroudakis G, Kotteas EA, Kotoula V, Papadopoulou K, Charalambous E, Cervantes A, et al. Mutational profiling of the RAS, PI3K, MET and b-catenin pathways in cancer of unknown primary: a retrospective study of the Hellenic Cooperative Oncology Group. Clinical & experimental metastasis. 2014;31(7):761-9.
29. Lu MY, Chen TY, Williamson DFK, Zhao M, Shady M, Lipkova J, et al. AI-based pathology predicts origins for cancers of unknown primary. Nature. 2021;594(7861):106-10.
30. Loffler H, Kramer, A. Biologic Features of CUP. In: Kramer A, Loffler, H, editor. Cancer of Unknown Primary Switzerland: Springer International Publishing; 2016.
31. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. 2018.
32. World Cancer Research Fund/American Institute for Cancer Research. Recommendations and public health and policy implications. 2018.
33. Anderson AS, Key TJ, Norat T, Scoccianti C, Cecchini M, Berrino F, et al. European code against cancer 4th edition: obesity, body fatness and cancer. Cancer Epidemiol 2015;39:S34-S45.
34. Leitzmann M, Powers H, Anderson AS, Scoccianti C, Berrino F, Boutron-Ruault M-C, et al. European Code against Cancer 4th Edition: Physical activity and cancer. Cancer Epidemiol. 2015;39:S46-S55.

35. Leon ME, Peruga A, McNeill A, Kralikova E, Guha N, Minozzi S, et al. European code against cancer: tobacco and cancer. Cancer Epidemiol 2015;39:S20-S33.
36. Norat T, Scoccianti C, Boutron-Ruault M-C, Anderson A, Berrino F, Cecchini M, et al. European code against cancer 4th edition: diet and cancer. Cancer Epidemiol. 2015;39:S56-S66.
37. Schüz J, Espina C, Villain P, Herrero R, Leon ME, Minozzi S, et al. European Code against Cancer 4th Edition: 12 ways to reduce your cancer risk. Cancer Epidemiol 2015;39:S1-S10.
38. Scoccianti C, Cecchini M, Anderson AS, Berrino F, Boutron-Ruault M-C, Espina C, et al. European Code against Cancer 4th Edition: Alcohol drinking and cancer. Cancer Epidemiol 2016;45:181-8.
39. Wild CP, Bucher JR, De Jong BW, Dillner J, Von Gertten C, Groopman JD, et al. Translational cancer research: balancing prevention and treatment to combat cancer globally. J Natl Cancer Inst. 2015;107(1):1-5.
40. Romaguera D, Vergnaud AC, Peeters PH, van Gils CH, Chan DS, Ferrari P, et al. Is concordance with World Cancer Research Fund/American Institute for Cancer Research guidelines for cancer prevention related to subsequent risk of cancer? Results from the EPIC study. Am J Clin Nutr. 2012;96(1):150-63.
41. Vergnaud AC, Romaguera D, Peeters PH, van Gils CH, Chan DS, Romieu I, et al. Adherence to the World Cancer Research Fund/American Institute for Cancer Research guidelines and risk of death in Europe: results from the European Prospective Investigation into Nutrition and Cancer cohort study1,4. Am J Clin Nutr. 2013;97(5):1107-20.
42. Van den Brandt PA, Goldbohm RA, van 't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol. 1990;43(3):285-95.
43. Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, et al. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin Nutr. 1994;48(4):253-65.
44. Goldbohm RA, van 't Veer P, van den Brandt PA, van 't Hof MA, Brants HA, Sturmans F, et al. Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. Eur J Clin Nutr. 1995;49(6):420-9.
45. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol. 1990;19(3):553-8.

CHAPTER 2

ALCOHOL CONSUMPTION, CIGARETTE SMOKING AND CANCER OF UNKNOWN PRIMARY RISK

Karlijn E.P.E. Hermans, Piet A. van den Brandt,
Caroline Loef, Rob L. H. Jansen, and Leo J. Schouten

Int. J. Cancer. 2020; 1-12

Abstract

Cancer of Unknown Primary (CUP) is a metastasised malignancy with no identifiable primary tumour origin. Despite the frequent occurrence and bleak prognosis of CUP, research into its aetiology is scarce. This study investigates alcohol consumption, tobacco smoking and CUP risk. We used data from the Netherlands Cohort Study, a cohort that includes 120,852 participants aged 55-69 years, who completed a self-administered questionnaire on cancer risk factors at baseline. Cancer follow-up was established through record linkage to the Netherlands Cancer Registry and Dutch Pathology Registry. After 20.3 years of follow-up, 963 CUP cases and 4,288 subcohort members were available for case-cohort analyses. Multivariable adjusted hazard ratios (HRs) were calculated using proportional hazards models. In general, CUP risk increased with higher levels of alcohol intake (P_{trend} = 0.02). The association was more pronounced in participants who drank ≥30 grams of ethanol per day (HR: 1.57, 95% CI: 1.20-2.05) compared to abstainers. Current smokers were at an increased CUP risk (HR: 1.59, 95% CI: 1.29-1.97) compared to never smokers. We observed that the more cigarettes, or the longer a participant smoked, the higher the CUP risk was (P_{trend} = 0.003 and P_{trend} = 0.02, respectively). Interaction on additive scale was found for participants with the highest exposure categories of alcohol consumption and cigarette smoking frequency and CUP risk. Our findings demonstrate that alcohol consumption and cigarette smoking are associated with increased CUP risk. Lifestyle recommendations for cancer prevention regarding not drinking alcohol and avoiding exposure to smoking are therefore also valid for CUP.

Introduction

Cancer of Unknown Primary (CUP) is a heterogeneous group of metastasised malignancies with no identifiable primary tumour origin [1, 2]. Cancer treatment, if any, is generally based on the primary tumour origin, which makes treating CUP challenging. Another complexity, is the absence of consensus on a CUP definition. Due to the use of different definitions globally, it is difficult to compare this entity [3]. In the Netherlands, the cancer clinical practice guidelines advise to use the definition 'CUP' if the patient has a metastasis of an unknown primary tumour origin, based on a cytological and/or histological proven metastasis of a cancer [4]. In 2018, CUP accounted for approximately 1,300 incident cases in the country, this contributed to almost 2% of all cancers as registered by the Netherlands Cancer Registry (NCR) [5, 6].

CUP occurrence is equal in both sexes. On average, patients are aged 74 years at diagnosis [5]. The disease primarily concerns adenocarcinoma (ca. 60%) and undifferentiated carcinoma (ca. 20%), with the most common metastatic sites of presentation being the liver (ca. 40%) and lymph nodes (ca. 20%) [2, 7]. In the Netherlands, the overall median survival for patients with a CUP diagnosis between 2010 and 2012, was 1.7 months [2]. Despite the frequent occurrence and bleak prognosis of CUP, research into its aetiology is particularly scarce. Potential CUP risk factors that have previously been identified include diabetes, family history of cancer, waist circumference, and smoking [8-11].

Lifestyle recommendations for cancer prevention have described both alcohol consumption and tobacco smoking as modifiable cancer risk factors. These recommendations advise against drinking alcohol and avoiding exposure to smoking [12-15]. To date, few studies have investigated the association between alcohol consumption, tobacco smoking, and CUP risk [10, 11, 16]. Two prospective cohort studies investigated alcohol consumption and did not observe an association with CUP risk [10, 11]. Three studies demonstrated a strong association between smoking and CUP risk [10, 11, 16]. None of these studies, however, assessed the association between smoking duration and CUP risk, and one did not investigate smoking frequency [16]. Therefore, we aimed to investigate the association between alcohol consumption, tobacco smoking and the development of CUP in greater depth. We hypothesize that 1) CUP risk is higher in participants with a high intake of alcoholic drinks, 2) CUP risk is higher in participants who smoke, and that 3) there is a multiplicative or additive interaction effect between alcohol consumption, tobacco smoking and CUP risk.

Material and Methods

Design and study population

The Netherlands Cohort Study (NLCS) was started on 17 September 1986 and includes 120,852 participants aged 55-69 years at baseline from 204 Dutch municipalities. Data processing and analyses were based on the nested case-cohort design. Cases were derived from the full cohort while the number of person-years at risk for the full cohort was estimated from a subcohort of 5,000 participants who were randomly sampled from the full cohort at baseline [17].

Outcome measure

In this study, CUP is defined as a metastasised epithelial malignancy with no identifiable primary tumour origin after cytological and/or histological verification during a patient's lifetime. This CUP definition only includes epithelial malignancies (ICD-O-3: M-8000 - M-8570), which excludes for example sarcoma, lymphoma, mesothelioma, and melanoma.

Follow-up

Cancer follow-up was established through annual record linkage of the full cohort with the NCR and the Dutch Pathology Registry (PALGA), to identify CUP cases within the NLCS [18]. Information regarding the site of metastasis was obtained from the NCR, but data was only partially available and, therefore, supplementary information was requested and retrieved from PALGA pathology excerpts. These pathology excerpts were also used to determine whether cytological and/or histological confirmed cases had been correctly categorised in the data received from the NCR.

After 20.3 years of follow-up (17 September 1986 until 31 December 2006), data was available for a total of 1,353 potential CUP cases. After excluding those cases without microscopical confirmation or non-epithelial histology, a total of 1,073 CUP cases remained. These CUP cases were further subdivided: according to histology (adenocarcinoma, undifferentiated carcinoma, squamous cell carcinoma, neuroendocrine carcinoma, and other carcinoma); according to number of metastases (multiple metastases of the same type were counted as one metastatic site, e.g., bone metastases in hip and vertebra were counted as one); according to localisation of metastasis (up to four localisations); and according to survival duration (≤1 and >1 year after diagnosis). The subcohort consisted of 4,774

participants after excluding members who reported a history of cancer (except for skin cancer) at baseline. Participants were also excluded when there were missing values on alcohol consumption or cigarette smoking. As a result, 963 CUP cases and 4,288 subcohort members were available for investigation (see Figure 1).

Questionnaire data

All cohort members completed a self-administered questionnaire, which included detailed questions on alcohol consumption, tobacco smoking, and other cancer risk factors. Alcohol consumption was measured over the year preceding baseline, and was addressed by questions on beer, red wine, white wine, sherry, other fortified wines, liqueurs and liquor. Frequency of alcohol consumption ranged from 'never' to '6-7 times per week', and information on the number of glasses consumed per day. Participants who reported 'never' or consumed 'less than once per month' were considered abstainers. Four items from the questionnaire (red wine, white wine, sherry, and liqueurs) were combined into one single wine variable since these items were substantially correlated and separate analysis would have resulted in small numbers of subjects within each stratum. Mean daily alcohol consumption was calculated by using the computerized Dutch food composition table [19]. Based on pilot study data, standard glass sizes were defined as 200 ml for beer, 105 ml for wine, 80 ml for sherry, and 45 ml for both liqueurs and liquor, corresponding to 8, 10, 11, 7, and 13 grams of ethanol, respectively. The food frequency questionnaire was validated against a 9-day diet record. The Spearman correlation coefficient between alcohol consumption as assessed by the questionnaire and that estimated by the diet record was 0.89 for all subjects and 0.85 for alcohol consumers [20, 21]. Tobacco smoking was addressed through questions on baseline smoking status, and the ages at first exposure and last (if stopped) exposure to smoking. Questions were also asked about smoking frequency, and smoking duration (excluding stopping periods), for cigarette, cigar, and pipe smokers. As the vast majority of smoking subcohort members were cigarette smokers, analyses were restricted to that particular group. Based on the questionnaire data, the following cigarette smoking variables were constructed: cigarette smoking status (never, ex, current); frequency (cigarettes per day); duration (years); and time since smoking cessation (years). Participants who indicated that they had never smoked cigarettes were considered never smokers. Time since smoking cessation was calculated as age at baseline minus age at smoking cessation. To avoid collinearity problems, smoking frequency and smoking duration were centered as proposed by Leffondré et al. [22].

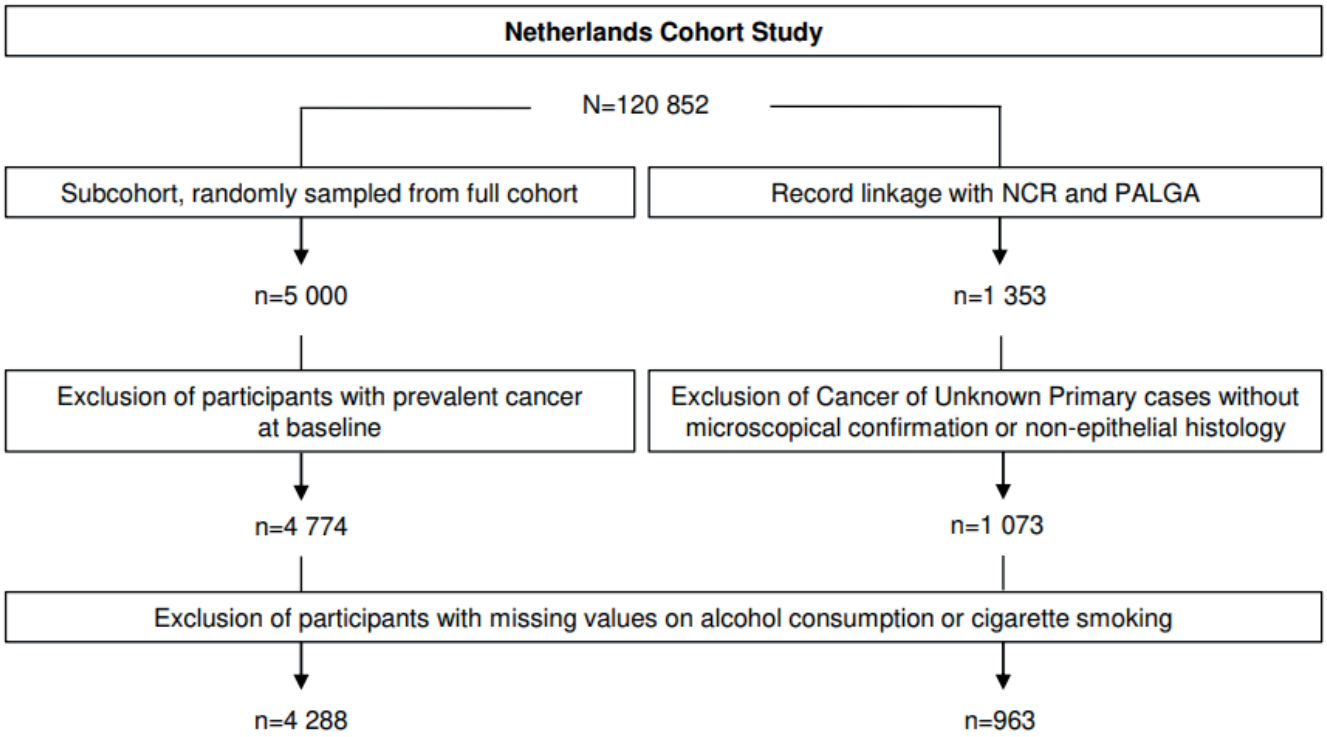


Figure 1 Flow diagram of subcohort members and Cancer of Unknown Primary cases in the Netherlands Cohort Study on whom analyses are based

Statistical analysis

Person-years at risk were calculated from baseline (17 September 1986) until CUP diagnosis, death, emigration, loss to follow-up, or end of follow-up (31 December 2006), whichever occurred first. Patient characteristics were presented for CUP cases and stratified for histological and cytological confirmation. General characteristics were presented for subcohort members and CUP cases with frequencies (percentages), for categorical variables and means, including standard deviations for continuous variables. Alcohol consumption was measured as a continuous variable with 10-gram ethanol per day increments and in categories: abstainers, >0-<5, 5-<15, 15-<30, and ≥30 grams of ethanol per day. Cigarette smoking was assessed based on status, frequency, duration, and time since smoking cessation. Cigarette smoking status was classified as never, ex, or current. Cigarette smoking frequency was measured as a continuous variable with 10 cigarettes smoked per day increments and in categories: never smokers, >0 to <10, 10-<20, and ≥20 cigarettes smoked per day. Cigarette smoking duration was investigated as a continuous variable with cigarette smoking increments of 10 years and in categories never smokers, >0-<20, 20-<40, and ≥40 years smoked. Time since smoking cessation was categorised as never smokers, stopped smoking ≥20 years, 10-<20 years, >0-<10 years, and current smokers. Alcohol consumption and cigarette smoking were mutually adjusted in the analyses. Predefined confounders were age at baseline (years; continuous), and sex (male/female). Potential confounders were body mass index (BMI) at baseline (kg/m^2), non-occupational physical activity (≤30 min/day, >30-60 min/day, >60-90 min/day and >90 min/day), socio-economic status (highest level of education), and history of cancer in a first degree relative (yes/no). Variables were considered a confounder if they are not an independent risk factor, not associated with the investigated exposure variables, and if the HR did not change by >10% when adding the potential confounder to the model. Accordingly, none of the potential confounders were included in the final model.

Cox proportional hazards models were used to evaluate associations of alcohol consumption, cigarette smoking, and CUP risk. Associations were estimated using age- and sex-adjusted, and multivariable adjusted hazard ratios (HRs) with 95% confidence intervals (CIs). Attributable risks were calculated for alcohol consumption and cigarette smoking. Standard errors were calculated using the robust Huber-White sandwich estimator to account for additional variance introduced by sampling from the full cohort [23]. The proportional hazards assumption was tested

using the scaled Schoenfeld residuals [24], and by visual inspection of log-minus-log (LML) survival curves. If there was an indication that the assumption had been violated, a time-varying covariate for that variable was added to investigate in the model. Tests for dose-response trends were assessed by fitting ordinal exposure variables as continuous terms. Wald tests and cross-product terms were used to evaluate potential interaction between alcohol consumption, sex, and CUP, between cigarette smoking, sex, and CUP, and multiplicative interaction between alcohol consumption, cigarette smoking frequency, and CUP. Interaction on additive scale between alcohol consumption, cigarette smoking frequency, and CUP was calculated using the relative excess risk due to interaction (RERI). Analyses were conducted using Stata version 15. *P* values were considered statistically significant if $p < 0.05$. Sensitivity analyses were conducted with restriction to histologically verified CUP cases, and after excluding the first two years and the first five years of follow-up to check for potential reverse causality bias.

Results

Data analysis was based on 963 CUP cases and 4,288 subcohort members for whom the data on alcohol consumption and cigarette smoking was complete. Overall, CUP patients were on average aged 73 years at diagnosis, the majority of whom were men (62.6%), and most cases were histologically verified (71.3%) (see Table 1). The most common histological subtype was adenocarcinoma (64.8%). The majority of patients were registered with a single organ metastasis (80.3%), and the most frequent metastatic site of presentation was the liver (37.9%). Most patients had died within a year after CUP diagnosis (73.4%).

Overall, CUP cases were more often alcohol consumers with a substantially higher ethanol intake (≥30 grams of ethanol) in comparison to subcohort members (16.6% versus 9.0%) (see Table 2). This higher intake was especially evident in male CUP patients of whom 23.9% drank ≥30 grams of ethanol per day on average in comparison to 14.7% of men in the subcohort. With respect to cigarette smoking, CUP cases were generally more often current smokers (37.8%) and less often never smokers (27.5%) in comparison to subcohort members (27.6% and 36.9%, respectively). Again, male CUP patients, in particular, were more often current smokers 44.9% in comparison to 34.8% of the men in the subcohort. In addition, the number of cigarettes smoked per day and smoking duration in years was higher for CUP patients on average in comparison to those of the subcohort members.

In general, a higher ethanol intake was associated with an increased CUP risk (P_{trend} = 0.02; see Table 3). Participants who reported consuming ≥30 grams of ethanol per day were compared to abstinence, at the highest risk of developing CUP (multivariable adjusted HR: 1.57, 95% CI: 1.20-2.05). The attributable risk for alcohol consumption on CUP risk was 4% (95% CI: 2%-6%). No multiplicative interaction was observed between alcohol consumption categories, sex, and CUP risk ($P_{interaction}$ = 0.86).

Current smokers were at an increased risk of developing CUP (multivariable adjusted HR: 1.59, 95% CI: 1.29-1.97) compared to never smokers (see Table 4). For cigarette smoking status, the attributable risk for CUP risk was 6% (95% CI: 4%-8%). After stratification for sex, we observed that CUP risk was the highest for current smokers compared to never smokers, in both men and women (HR: 1.64, 95% CI: 1.16-2.31 and HR: 1.62, 95% CI: 1.21-2.16, respectively). We observed that the more cigarettes a participant smoked, the higher the CUP risk (P_{trend} = 0.003). This trend was also observed in men (P_{trend} = 0.004). Overall, participants who smoked 10 to <20, or ≥20 cigarettes per day had a higher CUP risk (HR: 1.27, 95% CI: 1.00-1.62 and HR: 1.42, 95% CI: 1.13-1.80, respectively) compared to never smokers. There was no multiplicative interaction between cigarette smoking frequency, sex, and CUP risk ($P_{interaction}$ = 0.68). Additionally, we noted that the longer a participant had smoked cigarettes, the higher the CUP risk (P_{trend} = 0.02). Participants who smoked cigarettes ≥40 years were at the highest risk of developing CUP (HR: 1.45, 95% CI: 1.09-1.94) compared to never smokers. We found no multiplicative interaction between cigarette smoking duration, sex, and CUP risk ($P_{interaction}$ = 0.17). Categories of cigarette smoking cessation were assessed in comparison to never smokers. Participants who stopped <10 years were at a higher CUP risk (HR: 1.26, 95% CI: 0.99-1.62) compared to never smokers (P_{trend} <0.001). A similar trend was observed in men (P_{trend} <0.001) and in women (P_{trend} = 0.004).

Table 1 Patient characteristics of Cancer of Unknown Primary cases in the Netherlands Cohort Study

	Cancer of Unknown Primary cases					
	Overall (n=963)		Histologically confirmed (n=687)		Cytologically confirmed (n=276)	
	n	(%)	n	(%)	n	(%)
Age at baseline (years), %						
55-59	288	29.9	224	32.6	64	23.2
60-64	372	38.6	259	37.7	113	40.9
65-69	303	31.5	231	29.7	99	35.9
Age at diagnosis (years), mean (SD)						
Overall	73.3 (6.4)		72.8 (6.4)		74.7 (6.1)	
Sex, %						
Men	603	62.6	440	64.1	163	59.1
Women	360	37.4	247	36.0	113	40.9
Histology, %						
Adenocarcinoma	624	64.8	440	64.1	184	66.7
Undifferentiated carcinoma	192	19.9	133	19.4	59	21.4
Squamous cell carcinoma	47	4.9	38	5.5	9	3.3
Neuroendocrine carcinoma	35	3.6	32	4.7	3	1.1
Other carcinoma	65	6.8	44	6.4	21	7.6
Number of metastatic sites, %						
1	773	80.3	534	77.7	239	86.6
2+	166	17.2	140	20.4	26	9.4
Most frequent metastatic site of presentation, %						
Liver	365	37.9	307	44.7	58	21.0
Lymph node	158	16.4	118	17.2	40	14.5
Peritoneum	160	16.6	99	14.4	61	22.1
Bone	149	15.5	125	18.2	24	8.7
Lung	78	8.1	43	6.3	35	12.7
Survival status, %						
Survival ≤1 year after diagnosis	707	73.4	496	72.2	211	76.5
Survival >1 year after diagnosis	256	26.6	191	27.8	65	23.6

Table 2 Characteristics of Cancer of Unknown Primary cases and subcohort members in the Netherlands Cohort Study

	Subcohort members						Cancer of Unknown Primary cases					
	Total (M+F) (n=4288)		Men only (n=2110)		Women only (n=2178)		Total (M+F) (n=963)		Men only (n=603)		Women only (n=360)	
Exposure variables and potential confounders	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
Age at baseline (years)												
55-59	1664	38.8	818	38.8	846	38.8	288	29.9	187	31.0	101	28.1
60-64	1461	34.1	730	34.6	731	33.6	372	38.6	231	38.3	141	39.2
65-69	1163	27.1	562	26.6	601	27.6	303	31.5	185	30.7	118	32.8
Sex	4288	100	2110	49.2	2178	50.8	963	100	603	62.6	360	37.4
Alcohol consumption												
Ethanol intake (grams/day)												
Abstainers	1024	23.9	313	14.8	711	32.6	186	19.3	71	11.8	115	31.9
<5	1228	28.6	439	20.8	789	36.2	247	25.7	115	19.1	132	36.7
5-<15	979	22.8	576	27.3	403	18.5	217	22.5	146	24.2	71	19.7
15-<30	672	15.7	471	22.3	201	9.2	153	15.9	127	21.1	26	7.2
≥30	385	9.0	311	14.7	74	3.4	160	16.6	144	23.9	16	4.4
Ethanol intake (10 g ethanol/day increments), mean (SD) [1]	2.5 (1.5)		2.7 (1.7)		2.1 (1.1)		3.0 (1.9)		3.2 (1.9)		2.2 (1.2)	
Cigarette smoking												
Cigarette smoking status												
Never smokers	1584	36.9	291	13.8	1293	59.4	265	27.5	61	10.1	204	56.7
Ex smokers	1521	35.5	1085	51.4	436	20.0	334	34.7	271	44.9	63	17.5
Current smokers	1183	27.6	734	34.8	449	20.6	364	37.8	271	44.9	93	25.8
Frequency of cigarette smoking (N/day), mean (SD) [2]	15.7 (10.1)		17.3 (10.5)		12.3 (8.1)		17.8 (10.4)		19.0 (10.7)		13.4 (7.8)	
Duration of cigarette smoking (years), mean (SD) [2]	31.9 (12.1)		33.8 (11.6)		28.0 (12.2)		35.6 (11.6)		37.0 (10.8)		30.7 (12.7)	
Cigarette smoking cessation (years), mean (SD) [2]	14.5 (9.7)		14.8 (9.4)		13.8 (10.5)		12.9 (9.2)		12.7 (9.0)		13.3 (10.2)	

	Subcohort members						Cancer of Unknown Primary cases					
	Total (M+F) (n=4288)		Men only (n=2110)		Women only (n=2178)		Total (M+F) (n=963)		Men only (n=603)		Women only (n=360)	
Exposure variables and potential confounders	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
Other risk factors												
Body Mass Index at baseline (kg/m²), mean (SD)	25.0 (3.1)		25.0 (2.6)		25.1 (3.5)		25.0 (3.0)		24.9 (2.7)		25.1 (3.5)	
Non-occupational physical activity (min/day)												
≤30	908	21.5	390	18.7	518	24.2	204	21.5	108	18.1	96	27.1
>30-60	1318	31.2	633	30.4	685	31.9	291	30.6	181	30.4	110	31.1
>60-90	879	20.8	396	19.0	483	22.5	170	17.9	93	15.6	77	21.8
>90	1122	26.5	663	31.8	459	21.4	285	30.0	214	35.9	71	20.1
Level of education (years of education)												
Primary	1257	29.5	535	25.5	722	33.3	271	28.5	139	23.2	132	37.3
Lower vocational	937	22.0	429	20.4	508	23.4	204	21.4	117	19.6	87	24.6
Secondary and medium vocational	1483	34.8	739	35.2	744	34.3	341	35.8	234	39.1	107	30.2
University and higher vocational	590	13.8	396	18.9	194	9.0	136	14.3	108	18.1	28	7.9
History of cancer in a first degree relative												
Yes	1694	45.2	821	43.3	873	47.2	410	48.5	264	48.3	146	48.8

Notes
[1] In consumers only
[2] In users only

Table 3 Hazard ratios and 95% confidence intervals for alcohol consumption and Cancer of Unknown Primary risk in the Netherlands Cohort Study

		Subcohort members		Cancer of Unknown Primary cases					
		Categorical median	Person time at risk (years)	Cases	Age and sex-adjusted [1]		Multivariable adjusted [2]		
				n	HR	95% CI	HR	95% CI	
Ethanol intake (grams/day)									
Overall									
	Abstainers	0	17 180	186	1	Reference	1	Reference	
	>0-<5	2	21 482	247	1.06	(0.86- 1.31)	1.10	(0.88-	1.36)
	5-<15	9	16 214	217	1.11	(0.88- 1.38)	1.13	(0.90-	1.41)
	15-<30	22	11 159	153	1.05	(0.82- 1.34)	0.97	(0.76-	1.25)
	≥30	40	6 307	160	1.82	(1.41- 2.36)	1.57	(1.20-	2.05)
	p for trend [3]				<0.001		0.02		
	Continuous, 10 grams ethanol/day increments		72 342	963	1.12	(1.07- 1.17)	1.08	(1.03-	1.13)
Men									
	Abstainers	0	4 755	71	1	Reference	1	Reference	
	>0-<5	2	7 051	115	1.16	(0.82- 1.62)	1.19	(0.84-	1.67)
	5-<15	9	9 057	146	1.17	(0.85- 1.62)	1.20	(0.86-	1.67)
	15-<30	22	7 637	127	1.18	(0.85- 1.64)	1.08	(0.77-	1.51)
	≥30	41	5 029	144	2.00	(1.43- 2.79)	1.71	(1.21-	2.41)
	p for trend [3]				<0.001		0.01		
	Continuous, 10 grams ethanol/day increments		33 528	603	1.13	(1.08- 1.18)	1.09	(1.03-	1.14)
Women									
	Abstainers	0	12 425	115	1	Reference	1	Reference	
	>0-<5	2	14 431	132	1.00	(0.76- 1.31)	1.03	(0.78-	1.35)
	5-<15	9	7 157	71	1.09	(0.79- 1.51)	1.09	(0.79-	1.51)
	15-<30	21	3 522	26	0.83	(0.53- 1.31)	0.80	(0.50-	1.27)
	≥30	37	1 279	16	1.51	(0.84- 2.72)	1.28	(0.70-	2.35)
	p for trend [3]				0.63		0.92		
	Continuous, 10 grams ethanol/day increments		38 814	360	1.04	(0.91- 1.18)	1.00	(0.88-	1.14)

Notes

[1] Analyses were adjusted for age at baseline (years), and sex.

[2] Multivariable analyses were additionally adjusted for cigarette smoking status (never/former/current), frequency (continuous; centered), and duration (continuous; centered).

[3] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

Table 4A Hazard ratios and 95% confidence intervals for cigarette smoking and Cancer of Unknown Primary risk in the Netherlands Cohort Study

	Total (M+F)						
	Subcohort members (n=4288)		Cancer of Unknown Primary cases (n=963)				
	Categorical median	Person time at risk (years)	Cases	Age and sex-adjusted [1]		Multivariable adjusted [2]	
			n	HR	95% CI	HR	95% CI
Cigarette smoking status [4]							
Never smokers		28 472	265	1	Reference	1	Reference
Ex smokers		25 427	334	1.09	(0.90- 1.33)	1.19	(0.97- 1.47)
Current smokers		18 443	364	1.85	(1.53- 2.24)	1.59	(1.29- 1.97)
p for trend [3]				<0.001		<0.001	
Cigarette smoking frequency (N/day) [5]							
Never smokers	0	28 472	265	1	Reference	1	Reference
>0 to <10	5	12 521	119	0.94	(0.74- 1.19)	0.86	(0.65- 1.14)
10 to <20	12	15 252	264	1.59	(1.28- 1.97)	1.27	(1.00- 1.62)
≥20	20	16 097	315	1.83	(1.48- 2.25)	1.42	(1.13- 1.80)
p for trend [3]				<0.001		0.003	
Continuous, 10 cigarettes/day increments		72 342	963	1.21	(1.12- 1.30)	1.17	(1.08- 1.27)
Cigarette smoking duration (years) [6]							
Never smokers	0	28 472	265	1	Reference	1	Reference
>0 to <20	13	8 435	82	0.96	(0.73- 1.28)	0.95	(0.71- 1.27)
20 to <40	30	21 338	279	1.24	(1.02- 1.51)	1.07	(0.86- 1.33)
≥40	43	14 097	337	2.03	(1.65- 2.49)	1.45	(1.09- 1.94)
p for trend [3]				<0.001		0.02	
Continuous, 10 year increments		72 342	963	1.35	(1.24- 1.47)	1.18	(1.07- 1.30)

		Total (M+F)						
		Subcohort members (n=4288)		Cancer of Unknown Primary cases (n=963)				
		Categorical median	Person time at risk (years)	Cases	Age and sex-adjusted [1]		Multivariable adjusted [2]	
				n	HR	95% CI	HR	95% CI
Time since cigarette smoking cessation (years) [7]								
	Never smokers		28 472	265	1	Reference	1	Reference
	Stopped ≥20 years	25	7 817	79	0.80	(0.59- 1.07)	0.91	(0.67- 1.23)
	Stopped 10 to <20 years	14	8 678	110	1.05	(0.81- 1.36)	1.06	(0.81- 1.38)
	Stopped >0 to <10 years	5	8 861	144	1.40	(1.10- 1.78)	1.26	(0.99- 1.62)
	Current smokers	0	18 443	364	1.85	(1.52- 2.23)	1.67	(1.37- 2.03)
	p for trend [3]				<0.001		<0.001	

Notes

[1] Analyses were adjusted for age at baseline (years) and sex.

[2] Multivariable analyses were adjusted for age at baseline (years), sex, and alcohol consumption (g ethanol/day).

[3] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

[4] Multivariable analyses of cigarette smoking status were additionally adjusted for frequency (N/day; continuous; centered) and duration of cigarette smoking (years; continuous; centered).

[5] Multivariable analyses of cigarette smoking frequency were additionally adjusted for current cigarette smoking and duration of cigarette smoking (years; continuous; centered).

[6] Multivariable analyses of cigarette smoking duration were additionally adjusted for current cigarette smoking and frequency of cigarette smoking (N/day; continuous; centered).

[7] Cigarette smoking cessation was additionally adjusted for the number of cigarette pack-years (continuous; centered).

Table 4B Hazard ratios and 95% confidence intervals for cigarette smoking and Cancer of Unknown Primary risk in the Netherlands Cohort Study

	Men only								
	Subcohort members (n=2110)		Cancer of Unknown Primary cases (n=603)						
	Categorical median	Person time at risk (years)	Cases	Age and sex-adjusted [1]			Multivariable adjusted [2]		
			n	HR	95% CI		HR	95% CI	
Cigarette smoking status [4]									
Never smokers		5 026	61	1	Reference		1	Reference	
Ex smokers		17 558	271	1.30	(0.95-	1.78)	1.24	(0.90-	1.70)
Current smokers		10 945	271	2.27	(1.66-	3.11)	1.64	(1.16-	2.31)
p for trend [3]				<0.001			0.004		
Cigarette smoking frequency (N/day) [5]									
Never smokers	0	5 026	61	1	Reference		1	Reference	
>0 to <10	5	5 488	68	1.02	(0.69-	1.50)	0.91	(0.60-	1.37)
10 to <20	12	10 622	205	1.64	(1.19-	2.27)	1.27	(0.90-	1.78)
≥20	20	12 392	269	1.98	(1.44-	2.70)	1.49	(1.07-	2.07)
p for trend [3]				<0.001			0.004		
Continuous, 10 cigarettes/day increments		33 528	603	1.21	(1.12-	1.31)	1.17	(1.07-	1.28)
Cigarette smoking duration (years) [6]									
Never smokers	0	5 026	61	1	Reference		1	Reference	
>0 to <20	14	4 398	45	0.91	(0.59-	1.39)	0.86	(0.56-	1.33)
20 to <40	30	13 333	215	1.42	(1.03-	1.95)	1.20	(0.87-	1.66)
≥40	44	10 772	282	2.20	(1.61-	3.01)	1.49	(1.02-	2.19)
p for trend [3]				<0.001			0.01		
Continuous, 10 year increments		33 528	603	1.40	(1.26-	1.54)	1.23	(1.08-	1.39)

	Men only								
	Subcohort members (n=2110)		Cancer of Unknown Primary cases (n=603)						
	Categorical median	Person time at risk (years)	Cases	Age and sex-adjusted [1]			Multivariable adjusted [2]		
			n	HR	95% CI		HR	95% CI	
Time since cigarette smoking cessation (years) [7]									
Never smokers		5 026	61	1	Reference		1	Reference	
Stopped ≥20 years	25	5 738	60	0.86	(0.58-	1.28)	0.90	(0.60-	1.34)
Stopped 10 to <20 years	14	6 057	94	1.31	(0.91-	1.88)	1.21	(0.84-	1.74)
Stopped >0 to <10 years	5	5 754	117	1.76	(1.24-	2.51)	1.44	(1.00-	2.07)
Current smokers	0	10 945	271	2.28	(1.66-	3.12)	1.88	(1.36-	2.59)
p for trend [3]				<0.001			<0.001		

Notes

[1] Analyses were adjusted for age at baseline (years) and sex.

[2] Multivariable analyses were adjusted for age at baseline (years), sex, and alcohol consumption (g ethanol/day).

[3] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

[4] Multivariable analyses of cigarette smoking status were additionally adjusted for frequency (N/day; continuous; centered) and duration of cigarette smoking (years; continuous; centered).

[5] Multivariable analyses of cigarette smoking frequency were additionally adjusted for current cigarette smoking and duration of cigarette smoking (years; continuous; centered).

[6] Multivariable analyses of cigarette smoking duration were additionally adjusted for current cigarette smoking and frequency of cigarette smoking (N/day; continuous; centered).

[7] Cigarette smoking cessation was additionally adjusted for the number of cigarette pack-years (continuous; centered).

Table 4C Hazard ratios and 95% confidence intervals for cigarette smoking and Cancer of Unknown Primary risk in the Netherlands Cohort Study

	Women only							
	Subcohort members (n=2178)		Cancer of Unknown Primary cases (n=360)					
	Categorical median	Person time at risk (years)	Cases	Age and sex-adjusted [1]		Multivariable adjusted [2]		
			n	HR	95% CI	HR	95% CI	
Cigarette smoking status [4]								
Never smokers		23 446	204	1	Reference	1	Reference	
Ex smokers		7 869	63	0.99	(0.73- 1.34)	1.21	(0.87- 1.68)	
Current smokers		7 498	93	1.59	(1.21- 2.09)	1.62	(1.21- 2.16)	
p for trend [3]				0.003		0.001		
Cigarette smoking frequency (N/day) [5]								
Never smokers	0	23 446	204	1	Reference	1	Reference	
>0 to <10	4	7 033	51	0.88	(0.64- 1.23)	0.86	(0.56- 1.30)	
10 to <20	12	4 630	59	1.67	(1.21- 2.31)	1.45	(0.96- 2.19)	
≥20	20	3 704	46	1.60	(1.12- 2.28)	1.34	(0.88- 2.03)	
p for trend [3]				0.001		0.23		
Continuous, 10 cigarettes/day increments		38 814	360	1.22	(1.00- 1.49)	1.19	(0.97- 1.45)	
Cigarette smoking duration (years) [6]								
Never smokers	0	23 446	204	1	Reference	1	Reference	
>0 to <20	11	4 037	37	1.14	(0.78- 1.67)	1.27	(0.83- 1.96)	
20 to <40	30	8 005	64	1.03	(0.76- 1.39)	0.93	(0.65- 1.35)	
≥40	41	3 325	55	1.97	(1.40- 2.76)	1.56	(0.96- 2.54)	
p for trend [3]				0.004		0.27		
Continuous, 10 year increments		38 814	360	1.22	(1.03- 1.45)	1.09	(0.92- 1.29)	

	Women only								
	Subcohort members (n=2178)		Cancer of Unknown Primary cases (n=360)						
	Categorical median	Person time at risk (years)	Cases	Age and sex-adjusted [1]			Multivariable adjusted [2]		
			n	HR	95% CI		HR	95% CI	
Time since cigarette smoking cessation (years) [7]									
Never smokers		23 446	204	1	Reference		1	Reference	
Stopped ≥20 years	27	2 079	19	1.10	(0.66-	1.83)	1.42	(0.82-	2.48)
Stopped 10 to <20 years	14	2 621	16	0.77	(0.45-	1.33)	0.88	(0.51-	1.51)
Stopped >0 to <10 years	4	3 108	27	1.08	(0.70-	1.66)	1.13	(0.73-	1.74)
Current smokers	0	7 498	93	1.59	(1.21-	2.09)	1.59	(1.20-	2.11)
p for trend [3]				0.005			0.004		

Notes

[1] Analyses were adjusted for age at baseline (years) and sex.

[2] Multivariable analyses were adjusted for age at baseline (years), sex, and alcohol consumption (g ethanol/day).

[3] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

[4] Multivariable analyses of cigarette smoking status were additionally adjusted for frequency (N/day; continuous; centered) and duration of cigarette smoking (years; continuous; centered).

[5] Multivariable analyses of cigarette smoking frequency were additionally adjusted for current cigarette smoking and duration of cigarette smoking (years; continuous; centered).

[6] Multivariable analyses of cigarette smoking duration were additionally adjusted for current cigarette smoking and frequency of cigarette smoking (N/day; continuous; centered).

[7] Cigarette smoking cessation was additionally adjusted for the number of cigarette pack-years (continuous; centered).

We observed no multiplicative interaction between alcohol consumption, cigarette smoking frequency, and CUP risk ($P_{interaction}$ = 0.12) (see Table 5). However, we did find increased risks for most exposure combinations of alcohol consumption and cigarette smoking categories, for participants who smoked 10-<20, or ≥20 cigarettes per day compared to abstainers and never smokers as the reference group. In addition, we found that participants who drank ≥30 grams of ethanol per day and who smoked ≥20 cigarettes per day had the highest risk of developing CUP (HR: 2.87, 95% CI: 1.95-4.22) compared to abstainers and never smokers. We also assessed whether there was interaction on additive scale between the highest exposure categories of alcohol consumption (≥30 grams of ethanol per day), cigarette smoking (smoking ≥10 cigarettes per day) and CUP risk in comparison to the lowest exposure categories of alcohol consumption (<30 grams of ethanol per day) and cigarette smoking (smoking <10 cigarettes per day). The RERI was 1.14 (95% CI: 0.33-1.96); *P* = 0.006, which indicates that there is interaction on additive scale (see Table 6).

Results from the sensitivity analysis with restriction to histologically verified CUP cases, and results after excluding the first two years and the first five years of follow-up did not differ substantially from the findings of the complete multivariable analysis (data not shown).

Discussion

In this prospective cohort study, alcohol consumption and cigarette smoking were found to be associated with CUP risk. Associations were increased for participants who drank ≥30 grams of ethanol per day. Current smokers were at an increased risk of developing CUP. The more cigarettes (N/day) and the longer (years) participants had smoked, the greater their CUP risk. No multiplicative interaction was observed between alcohol consumption, cigarette smoking frequency and CUP risk. We did, however, find an interaction on additive scale between the highest exposure categories of alcohol consumption and cigarette smoking frequency and CUP risk.

Hitherto, only two prospective cohort studies have investigated the association between alcohol consumption and CUP risk. The European study, the European Prospective Investigation into Cancer and Nutrition (EPIC), includes 651 incident CUP cases. Results from the cohort revealed an increased CUP risk (HR: 1.42) for patients who consumed >60 grams of ethanol per day compared to an intake of 0-12

g/day, although not statistically significant [10]. The Australian study is a prospective cohort study that compared 327 incident cancer registry-notified CUP cases to two sets of controls that were randomly selected (3:1) using incidence density sampling with replacement. They observed no associations between alcohol consumption and CUP risk compared to the metastatic cancer controls, and compared to the general cohort population controls [11]. In the NLCS, we have found a positive association between alcohol consumption and CUP risk. The association was more pronounced in participants who drank ≥30 grams of ethanol per day compared to abstainers. Additionally, our stratified analysis indicates that the CUP risk was especially increased in men. However, it should be noted that alcohol consumption categories differed between the European study, the Australian study and the NLCS, which makes it difficult to compare outcomes.

In an additional analysis in the European study, squamous cell carcinoma cases were deliberately excluded when assessing the association between alcohol consumption and CUP risk, because the majority of these cases had metastases in cervical lymph nodes, which could indicate the primary origin to be a tumour in the upper aerodigestive tract [10]. After excluding squamous cell carcinoma cases from our cohort (N=47), no notable changes were identified for the association between alcohol consumption and CUP risk (data not shown).

The European study demonstrated that current heavy smokers (26+ cigarettes/day) had an increased risk of developing CUP (HR: 3.66) compared to never smokers [10]. Similarly, the Australian study reported that current smokers (odds ratio (OR): 3.42), or ex-smokers (OR: 1.95) were associated with CUP risk compared to never smokers [11]. A Swedish case-control study used data on 463 CUP patients, this study indicated that smoking was a risk factor for CUP (OR: 1.82) compared to no smoking [16]. However, the exposure category of no smoking was not described in-depth, and possibly included ex-smokers. In the NLCS, we also found current cigarette smokers to be at a greater risk of developing CUP (HR: 1.59) compared to never smokers. Although this association between smoking and CUP is weaker compared to those findings in the abovementioned studies, it should be noted that our study used different categories for measuring cigarette smoking. However, in accordance with the European study, we observed an association between smoking and CUP risk, which was elevated in the highest category of smoking frequency [10]. In contrast, the Australian study observed no difference in risk associated with <20, or ≥20 cigarettes per day [11].

Table 5 Interaction of alcohol consumption, cigarette smoking, and Cancer of Unknown Primary (CUP) risk (multivariable [1] adjusted incidence HRs) on multiplicative scale

	Alcohol consumption (grams/day)				
	Abstainers	**>0-<5**	**5-<15**	**15-<30**	**≥30**
Cigarette smoking frequency (N/day)					
Never smokers					
Cases	75	122	46	16	6
Person time at risk (years)	10 187	11 067	4 802	1 717	700
HR	1	1.52	1.20	1.17	0.96
95% CI	Reference	(1.11-2.08)	(0.80-1.80)	(0.65-2.09)	(0.39-2.38)
>0-<10					
Cases	13	32	40	21	13
Person time at risk (years)	2 137	3 742	3 559	2 219	864
HR	0.73	1.05	1.27	1.05	1.53
95% CI	(0.38-1.41)	(0.66-1.69)	(0.81-1.99)	(0.61-1.80)	(0.76-3.08)
10-<20					
Cases	49	47	67	58	43
Person time at risk (years)	2 522	3 653	4 348	3 151	1 578
HR	1.87	1.25	1.61	1.65	2.15
95% CI	(1.19-2.92)	(0.81-1.93)	(1.09-2.37)	(1.08-2.52)	(1.32-3.50)
≥20					
Cases	49	46	64	58	98
Person time at risk (years)	2 335	3 020	3 505	4 072	3 165
HR	2.05	1.66	1.82	1.32	2.87
95% CI	(1.32-3.18)	(1.09-2.54)	(1.22-2.73)	(0.87-2.01)	(1.95-4.22)
p for interaction [2]					0.115

Notes

[1] Adjusted for age at baseline (years), sex, current cigarette smoking, and duration of cigarette smoking (years; continuous; centered).

[2] *P* Value for interaction between categories of alcohol consumption and cigarette smoking, based on the Wald test and cross-product terms in the Cox proportional hazards model.

Table 6 Interaction of alcohol consumption, cigarette smoking, and Cancer of Unknown Primary (CUP) risk on additive scale

	Alcohol consumption	
	Low consumption (<30 grams of ethanol)	**High consumption (≥30 grams of ethanol)**
Cigarette smoking frequency		
Low exposure (smoking <10 cigarettes per day)	1 (reference)	1.82 (1.59-2.10)
High exposure (smoking ≥10 cigarettes per day)	1.31 (0.83-2.08)	3.28 (2.70-3.99)

Measure of interaction on additive scale: RERI = 1.14 (95% CI: 0.33-1.96); *P* = 0.006.

The European study also reported that participants who had quit smoking ≤10 years ago, were at a higher risk of CUP (HR: 1.34) than participants who had never smoked [10]. In the NLCS, we found that participants who had stopped <10 years were at a higher risk of developing CUP (HR: 1.26) compared to never smokers. Accordingly, our results are similar to those of the European study, which means our results are again in line with those of the European study.

Our study provides novel information on the association between cigarette smoking frequency, cigarette smoking duration and CUP risk. We found CUP risk to be more pronounced in the highest exposure categories of both cigarette smoking frequency and cigarette smoking duration, suggesting that the more cigarettes (N/day), or the longer (years), participants smoked, the greater their risk of developing CUP.

We found no multiplicative interaction effect between alcohol consumption, cigarette smoking frequency and CUP risk. However, we did find that participants who consumed the highest intake level of alcohol and smoked the highest number of cigarettes had a greater risk of CUP than either abstainers or never smokers. In addition, we found a significant additive interaction between the highest exposure categories of alcohol consumption and cigarette smoking frequency and CUP risk. This means that the combined effect of alcohol consumption and cigarette smoking frequency is larger than the sum of the individual effects of both alcohol consumption and cigarette smoking frequency [25]. It should however be acknowledged that for assessing the interaction on additive scale, exposure categories were generated in a dichotomous manner.

Strengths and limitations

An important strength of this study is its prospective cohort design. A further strength is that the NLCS consists of a large cohort of 120,852 participants who were followed up for cancer incidence by the cancer registry in the Netherlands. Cases were registered by trained registry clerks who had access to the medical files and entered data by applying uniform coding rules. Furthermore, we were able to analyse 963 incident CUP cases, which is a much higher number of cases than previous studies have used to investigate CUP aetiology. It should, however, be acknowledged that the CUP definition used here might differ from that used in other countries because the criteria for defining 'CUP' are heterogeneous. CUP cases within this study were consistently registered by NCR registry clerks, for which data was retrieved from pathology and

clinical reports [26]. Within the NLCS, information on alcohol consumption and exposure to smoking were collected before the outcome, minimizing the effect of information bias. A potential limitation of the current study is that data on all exposure variables are self-reported, which may have resulted in bias due to misclassification. However, we expect this misclassification to be non-differential. Another potential limitation is that the questions regarding smoking behaviour are not validated. Even so, the questions included detailed categories which the participant could answer. Unfortunately, we do not have data to check which diagnostic methods were used for our CUP patients. Nevertheless, if we restrict our analysis purely to histologically verified CUP cases, for whom extended diagnostic methods are more likely, we find that the results do not differ greatly from the complete multivariable analyses. Accordingly, the findings from the complete multivariable analyses are representative of CUP cases with or without an extensive diagnostic work-up.

Conclusions

In this study, alcohol consumption and cigarette smoking were found to be associated with an increased CUP risk. These findings suggest that lifestyle recommendations on cancer prevention regarding not drinking alcohol and avoiding exposure to smoking are also valid for CUP.

References

1. Loffler H, Kramer A. Introduction. In: Loffler H, Kramer A. Cancer of Unknown Primary ed.: Springer International Publishing Switzerland, 2016.
2. Schroten-Loef C, Verhoeven R, de Hingh I, van de Wouw A, van Laarhoven H, Lemmens V. Unknown primary carcinoma in the Netherlands: decrease in incidence and survival times remain poor between 2000 and 2012. Eur. J. Cancer 2018;101:77-86.
3. Kok S, van Laarhoven H, Lemmens V, Schroten C. International variation in diagnostic and treatment guidelines for carcinoma unknown primary. Cancer 2014;1:2015.
4. Oncoline. Primaire tumor onbekend. In: Pathologie NVv, ed. Literatuurbespreking primaire tumor onbekend, 2012.
5. Comprehensive Cancer Centre the Netherlands. Incidentie Primaire Tumor Onbekend: Integraal kankercentrum Nederland, 2020.
6. Meijer L, De Peuter R, Van der Zwan JM, Loef C. Primaire Tumor Onbekend: Integraal kankercentrum Nederland, 2020.
7. Hemminki K, Bevier M, Hemminki A, Sundquist J. Survival in cancer of unknown primary site: population-based analysis by site and histology. Ann Oncol 2012;23:1854-63.
8. Hemminki K, Försti A, Sundquist K, Li X. Cancer of unknown primary is associated with diabetes. Eur J Cancer Prev. 2016;25:246-51.
9. Hemminki K, Ji J, Sundquist J, Shu X. Familial risks in cancer of unknown primary: tracking the primary sites. J Clin Oncol 2011;29:435-40.
10. Kaaks R, Sookthai D, Hemminki K, Kramer A, Boeing H, Wirfalt E, Weiderpass E, Overvad K, Tjonneland A, Olsen A, Peeters PH, Bueno-de-Mesquita HB, et al. Risk factors for cancers of unknown primary site: Results from the prospective EPIC cohort. Int J Cancer 2014;135:2475-81.
11. Vajdic C, Perez-Concha O, Dobbins T, Ward RL, Schaffer AL, van Leeuwen M, Rhee JJ, Laaksonen M, Craigen G, Pearson S-A. Demographic, social and lifestyle risk factors for cancer registry-notified cancer of unknown primary site (CUP). Cancer Epidemiol. 2019;60:156-61.
12. Schüz J, Espina C, Villain P, Herrero R, Leon ME, Minozzi S, Romieu I, Segnan N, Wardle J, Wiseman M. European Code against Cancer 4th Edition: 12 ways to reduce your cancer risk. Cancer Epidemiol. 2015;39:S1-S10.
13. World Cancer Research Fund. WCRF Cancer Prevention Recommendations, vol. 2016: Word Cancer Research Fund, 2016.
14. Leon ME, Peruga A, McNeill A, Kralikova E, Guha N, Minozzi S, Espina C, Schuez J. European code against cancer: tobacco and cancer. Cancer Epidemiol. 2015;39:S20-S33.
15. Scoccianti C, Cecchini M, Anderson AS, Berrino F, Boutron-Ruault M-C, Espina C, Key TJ, Leitzmann M, Norat T, Powers H. European Code against Cancer 4th Edition: Alcohol drinking and cancer. Cancer Epidemiol. 2016;45:181-88.
16. Hemminki K, Chen B, Melander O, Manjer J, Hallmans G, Hemminki A. Smoking and body mass index as risk factors for subtypes of cancer of unknown primary. Int J Cancer 2015;136:246-7.
17. Van den Brandt PA, Goldbohm RA, van 't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol 1990;43:285-95.

18. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol 1990;19:553-8.
19. NEVO-table. Dutch food composition table 1986–1987. Voorlichtingsbureau voor de voeding, The Hague, The Netherlands 1986.
20. Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, Hermus RJ. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin Nutr 1994;48:253-65.
21. Maasland DH, van den Brandt PA, Kremer B, Goldbohm RA, Schouten LJ. Alcohol consumption, cigarette smoking and the risk of subtypes of head-neck cancer: results from the Netherlands Cohort Study. BMC Cancer 2014;14:187.
22. Leffondre K, Abrahamowicz M, Siemiatycki J, Rachet B. Modeling smoking history: a comparison of different approaches. American journal of epidemiology 2002;156:813-23.
23. Barlow WE. Robust variance estimation for the case-cohort design. Biometrics 1994:1064-72.
24. Lin DY, Wei L-J. The robust inference for the Cox proportional hazards model. J Am Stat Assoc 1989;84:1074-78.
25. Knol MJ, VanderWeele TJ, Groenwold RH, Klungel OH, Rovers MM, Grobbee DE. Estimating measures of interaction on an additive scale for preventive exposures. Eur. J. Epidemiol. 2011;26:433-38.
26. Van der Sanden G, Coebergh J-W, Schouten L, Visser O, Van Leeuwen F. Cancer incidence in the Netherlands in 1989 and 1990: First results of the nationwide Netherlands Cancer Registry. Eur. J. Cancer 1995;31:1822-29.

CHAPTER 3

ANTHROPOMETRY, PHYSICAL ACTIVITY AND CANCER OF UNKNOWN PRIMARY RISK

Karlijn E.P.E. Hermans, Piet A. van den Brandt,
Caroline Loef, Rob L. H. Jansen, and Leo J. Schouten

Cancer Epidemiology: 2020; 69

Abstract

Background: Cancer of Unknown Primary (CUP) is a metastatic disease for which the primary tumour origin could not be identified during life. Few studies have investigated the risk factors associated with this disease. This study investigates anthropometry, physical activity and CUP risk.

Methods: Data is used from the Netherlands Cohort Study, which includes 120,852 participants aged 55-69 years. All cohort members completed a self-administered questionnaire on cancer risk factors at baseline in 1986. Cancer follow-up was established through record linkage to the Netherlands Cancer Registry and the Dutch Pathology Registry. After a follow-up of 20.3 years, 926 incident CUP cases and 4,099 subcohort members were available for case-cohort analyses. Proportional hazards models were used to compute multivariable adjusted hazard ratios (HRs).

Results: We found no associations between height, body mass index (BMI) at baseline, BMI at age 20 years, change in BMI since age 20 years, clothing size (trouser/skirt size), or non-occupational physical activity and CUP risk.

Conclusion: Our findings indicate that neither anthropometry nor physical activity are associated with the development of CUP.

Introduction

Cancer of Unknown Primary (CUP) is a metastasised malignancy for which the primary tumour origin could not be identified during life (1-4). In the Netherlands, CUP accounted for approximately 1,300 incident cases in 2018, which represented almost 2% of all new cancer diagnoses (5, 6). The disease has a bleak prognosis with a median survival of 1.7 months (2000-2012) (2, 7, 8).

CUP is a highly complex entity and little is known about the potential risk factors that contribute to its development. To date, those potential risk factors that have been identified include tobacco smoking, diabetes, and a family history of cancer (9-11). CUP risk also appears to be increased with higher intake levels of alcohol consumption (12).

Previous studies have investigated the relationship between body fatness, physical activity and the development of cancer (13-18). For example, an excess energy intake and/or low energy expenditure can lead to obesity and associated metabolic alterations (14). Similarly, adipose tissue is recognised as metabolically active and thus a source of adipose-tissue-derived hormones and cytokines. These metabolic alterations can consequently have effects on carcinogenesis (14-16). Greater body fatness has been associated with an increased risk for cancers of the oesophagus, pancreas, colorectum, endometrium, breast (post-menopausal), and kidney (14, 18). Likewise, low levels of physical activity have been associated with weight gain, and may therefore contribute to diseases associated with being overweight or obese (17, 19). However, the stimulation of immune functioning through regular moderate physical activity could reduce cancer risk independent from the effects of high body mass (19, 20). Physical activity has been associated with a reduced risk for cancers of the colon, breast, and endometrium (13, 17, 18). As a result, lifestyle recommendations for cancer prevention advise maintaining a healthy weight and being physically active (14, 17, 18).

Hitherto, only two studies appear to have investigated the association between anthropometry and CUP risk (10, 21). The results of the European Prospective Investigation into Cancer (EPIC) demonstrated that a high waist circumference was associated with an increased CUP risk. However, the investigators did not find an association between body mass index (BMI) and CUP risk (10). Comparably, an Australian prospective cohort study observed no association between BMI and CUP risk (21). In addition, its authors further assessed the association between physical

activity and CUP risk and found that moderate or vigorous physical activity and sedentary behaviour were not associated with CUP risk (21).

The aim of the current study was to investigate the association between anthropometry, physical activity and CUP risk in more detail by assessing various anthropometric variables such as height, BMI at baseline, BMI at age 20 years, change in BMI since age 20 years, clothing size (trouser/skirt size), and physical activity, in a large number of exposure categories. For this purpose, we used data from the Netherlands Cohort Study (NLCS) which includes a population with many CUP cases, and a consistent disease definition.

Material and methods

Design and study population

The NLCS was established on 17 September 1986 and included 120,852 participants aged 55-69 years from 204 Dutch municipalities. For efficiency reasons, data processing and analysis were based on the case-cohort design. Cases were derived from the full cohort while the number of person-years at risk was estimated from a subcohort of 5,000 participants who were randomly sampled from the full cohort at baseline (22). The institutional review boards of the Netherlands Organization for Applied Scientific Research TNO (Zeist) and Maastricht University (Maastricht) approved the execution of the NLCS. Participants agreed to be included into the cohort and follow-up by returning their completed questionnaires.

Outcome measure

In this study, CUP is defined as a metastasised epithelial malignancy with no identifiable primary tumour origin after cytological and/or histological verification during a patient's lifetime. This CUP definition only includes epithelial malignancies (ICD-O-3: M-8000 - M-8570), which excludes for example sarcoma, lymphoma, mesothelioma, and melanoma.

Follow-up

Cancer follow-up was established through annual record linkage with the Netherlands Cancer Registry (NCR) and the Dutch Pathology Registry (PALGA) (23). Information regarding the site of metastasis was obtained from the NCR, but

this data was only partially available and, therefore, supplementary information was retrieved from the pathology excerpts provided by PALGA. These pathology excerpts were also used to determine whether cytological and/or histological confirmed cases had been correctly categorised in the data received from the NCR.

After 20.3 years of follow-up (17 September 1986 until 31 December 2006), data was available for a total of 1,353 potential CUP cases, and a subcohort of 4,774 participants could be formed after removing those members who had reported a history of cancer (except for skin cancer) at baseline. After excluding CUP cases without microscopical confirmation or non-epithelial histology, a total of 1,073 CUP cases remained. These CUP cases were further subdivided: according to histology (adenocarcinoma, undifferentiated carcinoma, squamous cell carcinoma, neuroendocrine carcinoma, and other carcinoma); according to number of metastases (multiple metastases of the same type were counted as one metastatic site, e.g., bone metastases in hip and vertebra were counted as one); according to localisation of metastasis (up to four localisations), and according to survival duration (≤1 and >1 year after diagnosis). Participants were also excluded when there were missing values on BMI at baseline, physical activity, or selected confounders. As a result, 926 CUP cases and 4,099 subcohort members were available for analysis (see Figure 1).

Questionnaire data

All cohort members completed a self-administered questionnaire, which included detailed questions on height, weight, weight at age 20 years, physical activity, alcohol consumption, smoking habits, and other cancer risk factors. Height was asked in centimetres (cm), and weight at baseline and weight at age 20 in kilograms (kg). BMI at baseline was calculated using weight at baseline whereas BMI at age 20 years was calculated using weight at that age, both divided by height at baseline squared (kg/m^2). Change in BMI since age 20 years was calculated as BMI at baseline minus BMI at age 20 years. Clothing size was determined by asking trouser size for men and skirt size for women (24). Physical activity was measured by focused questioning regarding recreational physical activity and the physical activity involved in going to and from work (e.g., walking, cycling). The following questions were asked: ‘How many minutes do you spend on average per day walking or cycling to your work, going shopping, or walking your dog?’ and ‘How many hours of leisure time do you spend on average per week on the following

activities 1) gardening, 2) cycling, walking, and 3) sports/physical exercise?'. The reported times on both questions were summed into a total non-occupational physical activity value.

The baseline questionnaire also included questions related to alcohol consumption and tobacco smoking. Alcohol consumption was measured over the year preceding baseline, and was addressed by asking questions about the consumption of beer, red wine, white wine, sherry, other fortified wines, liqueurs and liquor. Frequency of alcohol consumption was measured in ranges from 'never' to '6-7 times per week', and information was requested on the number of glasses consumed per day. Participants who reported 'never' or consumed 'less than once per month' were considered abstainers. Mean daily alcohol consumption was calculated by using the computerized Dutch food composition table (25). Tobacco smoking was addressed through questions on baseline smoking status, and the ages at first exposure and last (if stopped) exposure to smoking. Questions were also asked about smoking frequency and smoking duration (excluding stopping periods), for cigarette, cigar, and pipe smokers. Participants who indicated that they had never smoked cigarettes were considered never smokers.

Statistical analysis

Person-years at risk were calculated from baseline (17 September 1986) until CUP diagnosis, death, emigration, loss to follow-up, or end of follow-up (31 December 2006), whichever occurred first. Patient characteristics were presented for CUP cases and stratified for histological and cytological confirmation. General characteristics were presented for subcohort members and CUP cases with frequencies (percentages) for categorical variables, and means including standard deviations for continuous variables.

Measurements of anthropometry are reported as height (cm), BMI at baseline, BMI at age 20 years, and change in BMI since age 20 years (kg/m^2), and clothing size (trouser size and skirt size for men and women, respectively). Height was analysed as a continuous variable with 5 cm increments and in categories, for men: <170 (reference category), 170-<175, 175-<180, 180-<185, ≥185 cm, and for women: <160 (reference category), 160-<165, 165-<170, 170-<175, ≥175 cm. BMI at baseline was classified as <20 (underweight), 20 to <25 (normal weight) (reference category), 25 to <30 (overweight), and ≥30 kg/m^2 (obese). BMI at age 20 years was classified as

<20, 20-<21.5 (reference category), 21.5-<23, 23-<25, and ≥25 kg/m^2. Change in BMI since age 20 years was classified as <0, 0-<4 (reference category), 4-<8, and ≥8 kg/m^2. The three aforementioned BMI variables were also measured on a continuous scale. Trouser size for men was categorised as <50, 50-51 (reference category), 52-53, 54-55, >56. Skirt size for women was categorised as <40, 42 (reference category), 44, 46-48, >50. Clothing size was shown to be correlated with waist circumference measurements in men and women in this Dutch population (r = 0.64, r = 0.71, respectively) (24). Physical activity was measured as a continuous variable with 30-minute increments and in categories, of ≤30, >30-60, >60-90, and >90 min/day.

Predefined confounders included: age at baseline (years; continuous); sex (male/female); alcohol consumption (ethanol intake in grams per day); cigarette smoking status (never, ever); frequency (number of cigarettes smoked per day); and duration (number of years smoking). Height was also adjusted for weight at baseline (kilograms; continuous). BMI at baseline, and BMI at age 20 years were additionally adjusted for physical activity (minutes/day; categories) in the analyses. Change in BMI since the age of 20 years was also adjusted for BMI at age 20 years (kg/m^2; continuous). Clothing size was similarly adjusted for physical activity (minutes/day; categories). Physical activity was additionally adjusted for BMI at baseline (kg/m^2; continuous). Potential confounders included: socio-economic status (highest level of education); diabetes (yes/no); and history of cancer in a first-degree relative (yes/no). Variables were considered a confounder if they changed the HR by >10%. None of the potential confounders were included in the final model.

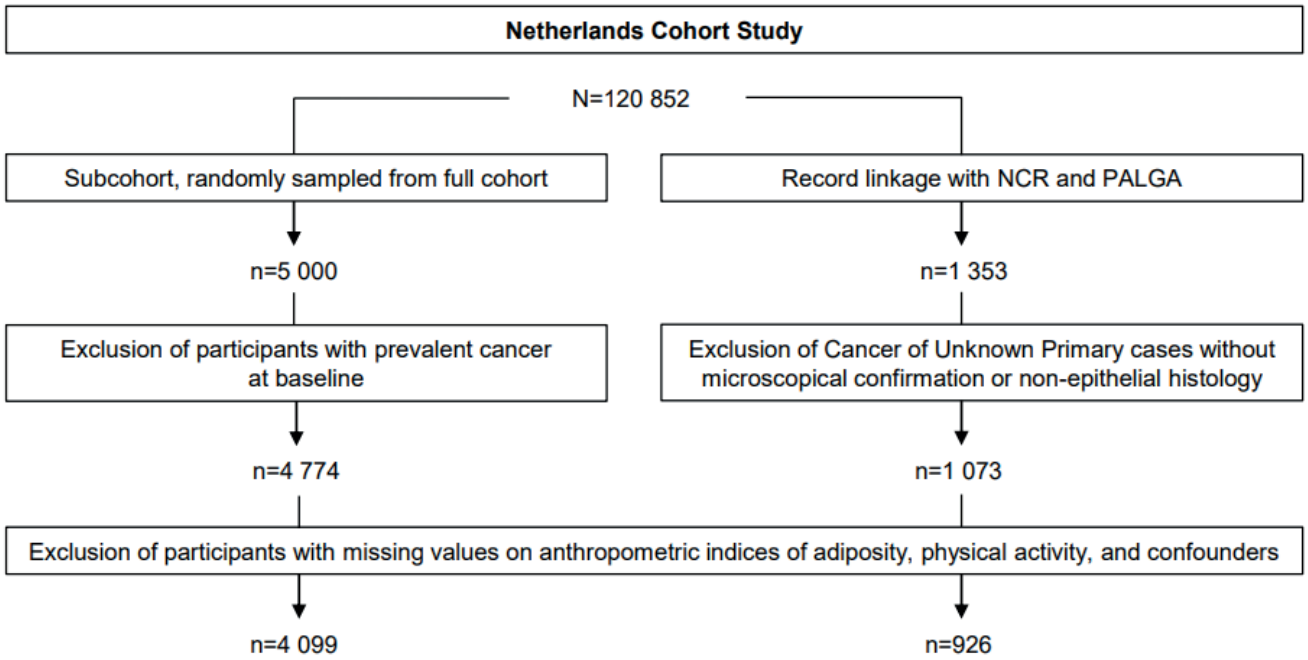


Figure 1 Flow diagram of subcohort members and Cancer of Unknown Primary cases in the Netherlands Cohort Study on whom analyses are based

Cox proportional hazards models were used to investigate associations of anthropometry, physical activity, and CUP risk. Associations were estimated using age- and sex-adjusted, and multivariable adjusted hazard ratios (HRs) with 95% confidence intervals (CIs). Standard errors were calculated using the robust Huber-White sandwich estimator to account for additional variance introduced by sampling from the full cohort (26). The proportional hazards assumption was tested using the scaled Schoenfeld residuals (27), and by visual inspection of log-minus-log (LML) survival curves. In case the assumption had been violated, a time-varying covariate (TVC) for that variable was added to the model if appropriate. The PH assumption appeared to be violated for the exposure variables change in BMI since the age of 20 years, and trouser size (men only). Inspection of the log-minus-log survival curves, however, revealed rather parallel survival curves and, hence, no violation of the PH assumption. Ordinal exposure variables were fitted as continuous variables in trend analyses. Wald tests and cross-product terms were used to evaluate potential multiplicative interaction between anthropometry and sex, between physical activity and sex, and to assess multiplicative interaction between BMI, physical activity, and CUP risk. If there was a significant interaction between exposure variables and sex, sex-stratified HRs were presented. Analyses were conducted using Stata version 15. *P* values were considered statistically significant if $p < 0.05$.

A sensitivity analysis was conducted with restriction to histologically verified CUP cases. It is more likely that patients with a histologically verified CUP underwent extensive diagnostic investigation(s) to rule out the primary tumour origin. For those patients who received cytological verification alone, other factors may have played a role in the decision to refrain from further diagnostic investigation such as age, comorbidities, performance status, localisation of metastasis, and most importantly the patient's decision. Another sensitivity analysis was performed after the first two years of follow-up had been excluded to check for potential reverse causality bias as a result of preclinical cancer at baseline. A third sensitivity analysis was performed to assess associations between anthropometry, physical activity, and CUP risk while taking the years of diagnosis into account by comparing the first 10 years of follow up (<1996) to the last 10 years of follow up (≥1996). It is possible that the casemix of CUP changed over time, as the definition of the CUP diagnosis has changed in the cancer registries throughout the years.

Results

Analyses are based on 926 incident CUP cases and 4,099 subcohort members for whom the data on BMI at baseline and physical activity was complete. Due to additional missing values with respect to BMI at age 20 years, 772 CUP cases and 3,516 subcohort members were available for that analysis.

CUP cases were, on average, aged 73 years at diagnosis, the majority of whom were men (62.5%), and most cases were histologically verified (71.1%) (see Table 1). The most common histological subtype was adenocarcinoma (64.8%). In the majority of cases, a single organ was affected by a metastasis (80.1%), and the most frequent metastatic site of presentation was the liver (38.0%). Most patients had died within a year after CUP diagnosis (73.1%).

Height differences were observed between CUP cases and subcohort members, but height was equally distributed in sex-stratified analyses (see Table 2). Male CUP cases were slightly lighter in comparison to male subcohort members (77.4 versus 77.9 kg, respectively). Female CUP cases and female subcohort members had a similar weight (68.5 versus 68.3 kg, respectively). The mean BMI at baseline was similar in CUP cases and subcohort members (25.0 and 25.0 kg/m^2, respectively). We observed only minor differences between CUP cases and subcohort members in the mean BMI at age 20 years (21.7 and 21.5 kg/m^2, respectively), as well as in change in BMI since age 20 years (3.3 and 3.5 kg/m^2, respectively). CUP cases were physically active for a longer amount of time compared to subcohort members (77.1 versus 73.0 min/day, respectively). Overall, a greater number of alcohol consumers with a substantially higher ethanol intake (≥30 grams of ethanol) was found in CUP cases in comparison to subcohort members (17.1% versus 9.1%, respectively). CUP cases were more frequently current smokers and had fewer never smokers in comparison to subcohort members (38.0% versus 27.8%, and 27.0% versus 36.5%, respectively). In addition, the number of cigarettes smoked per day and smoking duration in years was higher for CUP cases on average in comparison to those of the subcohort members (17.9 cigarettes per day versus 15.8 cigarettes per day, and 35.5 years of smoking versus 31.9 years of smoking, respectively). Diabetes was slightly more prevalent in CUP cases compared to subcohort members (3.8% versus 3.5%, respectively). More CUP cases had a history of cancer in a first-degree relative in comparison to subcohort members (47.6% versus 45.5%, respectively).

Results from the age- and sex-adjusted analyses (data not shown) were comparable to the results of the multivariable adjusted analyses. Therefore, we only reported the multivariable adjusted results. We observed a statistically significant multiplicative interaction between sex and the association between height and CUP risk ($P_{interaction}$ = 0.04). We found no association between height and CUP risk (HR in men: 0.98, 95% CI: 0.90-1.08, and HR in women: 1.04, 95% CI: 0.94-1.16) (see Table 3). BMI at baseline and BMI at age 20 years were not associated with CUP risk (HR: 0.99, 95% CI: 0.97-1.02, and HR: 1.03, 95% CI: 0.99-1.06, respectively) (see Table 4). We found no association between BMI change and CUP risk. Regarding the association between trouser size and CUP risk, no statistically significant association was observed (see Table 3). We found no statistically significant association for women who had a skirt size of >50 and CUP risk (HR: 1.51, 95% CI: 0.91-2.52), compared to women with a skirt size of 42. Overall, we observed no association between physical activity and CUP risk (HR: 1.01, 95% CI: 0.97-1.04) (see Table 5).

No multiplicative interaction was found between exposure combinations of BMI at baseline and physical activity categories and CUP risk ($P_{interaction}$ = 0.73) (data not shown). Results from the sensitivity analysis with restriction to histologically verified CUP cases did not differ substantially from the findings of the overall analyses. This was also true for both results after the first two years of follow-up had been excluded and for the results after evaluating the years of diagnosis with respect to the first 10 years of follow up in comparison to the last 10 years of follow up (data not shown).

Discussion

The current study has assessed various anthropometric variables, physical activity and their relation to CUP risk in a large prospective cohort study. The wide availability of those variables has allowed us to conduct multivariable analyses by investigating exposure categories in greater detail compared to those analyses performed in previous cohort studies. The findings presented here appear to demonstrate that anthropometry and physical activity are not associated with the development of CUP. We also found no multiplicative interaction between BMI, physical activity and CUP risk.

Table 1 Patient characteristics of Cancer of Unknown Primary cases in the Netherlands Cohort Study

	Cancer of Unknown Primary cases					
	Overall (n=926)		Histologically confirmed (n=658)		Cytologically confirmed (n=268)	
	n	(%)	n	(%)	n	(%)
Age at diagnosis (years), mean (SD)						
Overall	73.4 (6.4)		72.8 (6.4)		74.7 (6.1)	
Sex, %						
Men	579	62.5	422	64.1	157	58.6
Women	347	37.5	236	35.9	111	41.4
Histology, %						
Adenocarcinoma	600	64.8	421	63.4	179	66.8
Undifferentiated carcinoma	187	20.2	130	19.8	57	21.3
Squamous cell carcinoma	47	5.1	38	5.8	9	3.4
Neuroendocrine carcinoma	31	3.4	28	4.3	3	1.1
Other carcinoma	61	6.6	41	6.2	20	7.5
Number of metastatic sites, %						
1	742	80.1	511	77.7	231	86.2
2+	161	17.4	135	20.5	26	9.7
Most frequent metastatic site of presentation, %						
Liver	352	38.0	294	44.7	58	21.6
Lymph node	151	16.3	114	17.3	37	13.8
Peritoneum	155	16.7	95	14.4	60	22.4
Bone	143	15.4	120	18.2	23	8.6
Lung	75	8.1	43	6.5	32	11.9
Survival status, %						
Survival ≤1 year after diagnosis	677	73.1	473	71.9	204	76.1
Survival >1 year after diagnosis	249	26.9	185	28.1	64	23.9

Table 2 Characteristics of Cancer of Unknown Primary cases and subcohort members in the Netherlands Cohort Study

	Subcohort members						Cancer of Unknown Primary cases					
	Total (M+F) (n=4099)		Men only (n=2019)		Women only (n=2080)		Total (M+F) (n=926)		Men only (n=579)		Women only (n=347)	
Exposure variables and potential confounders	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
Age at baseline (years)												
55-59	1597	39.0	787	39.0	810	38.9	274	29.6	178	30.7	96	27.7
60-64	1401	34.2	698	34.6	703	33.8	360	38.9	222	38.3	138	39.8
65-69	1101	26.9	534	26.5	567	27.3	292	31.5	179	30.9	113	32.6
Sex	4099	100	2019	49.3	2080	50.7	926	100	579	62.5	347	37.5
Anthropometric variables												
Height (cm) at baseline, mean (SD)	170.8 (8.6)		176.6 (6.6)		165.2 (6.2)		172.1 (8.5)		176.1 (6.8)		165.4 (6.6)	
Weight (kg) at baseline, mean (SD)	73.1 (10.9)		77.9 (9.4)		68.3 (10.1)		74.1 (10.7)		77.4 (9.6)		68.5 (10.1)	
BMI (kg/m²) at baseline												
<20	144	3.5	41	2.0	103	5.0	26	2.8	10	1.7	16	4.6
20-<25	2082	50.8	1026	50.8	1056	50.8	487	52.6	313	54.1	174	50.1
25-<30	1614	39.4	879	43.5	735	35.3	352	38.0	227	39.2	125	36.0
≥30	259	6.3	73	3.6	186	8.9	61	6.6	29	5.0	32	9.2
BMI (kg/m²) at baseline, mean (SD)	25.0 (3.1)		25.0 (2.6)		25.1 (3.5)		25.0 (3.0)		24.9 (2.7)		25.0 (3.5)	
BMI (kg/m²) at age 20 years [1]												
<20	906	22.1	330	16.3	576	27.7	176	19.0	83	14.3	93	26.8
20-<21.5	840	20.5	427	21.2	413	19.9	197	21.3	136	23.5	61	17.6
21.5-<23	847	20.7	419	20.8	428	20.6	183	19.8	111	19.2	72	20.8
23-<25	650	15.9	324	16.1	326	15.7	151	16.3	98	16.9	53	15.3
≥25	273	6.7	128	6.3	145	7.0	65	7.0	33	5.7	32	9.2

	Subcohort members						Cancer of Unknown Primary cases					
	Total (M+F) (n=4099)		Men only (n=2019)		Women only (n=2080)		Total (M+F) (n=926)		Men only (n=579)		Women only (n=347)	
Exposure variables and potential confounders	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
BMI (kg/m²) at age 20 years, mean (SD) [1]	21.5 (2.6)		21.7 (2.4)		21.4 (2.7)		21.7 (2.7)		21.8 (2.4)		21.6 (3.0)	
Change in BMI (kg/m²) since age 20 years [1]												
<0	409	10.0	162	8.0	247	11.9	103	11.1	63	10.9	40	11.5
0-<4	1685	41.1	869	43.0	816	39.2	388	41.9	242	41.8	146	42.1
4-<8	1112	27.1	504	25.0	608	29.2	215	23.2	125	21.6	90	25.9
≥8	310	7.6	93	4.6	217	10.4	66	7.1	31	5.4	35	10.1
Change in BMI (kg/m²) since age 20 years, mean (SD) [1]	3.5 (3.3)		3.3 (2.9)		3.7 (3.6)		3.3 (3.5)		3.1 (3.2)		3.6 (3.7)	
Physical activity												
Non-occupational physical activity (min/day)												
≤30	861	21.0	371	18.4	490	23.6	199	21.5	105	18.1	94	27.1
>30-60	1276	31.1	613	30.4	663	31.9	285	30.8	177	30.6	108	31.1
>60-90	863	21.1	388	19.2	475	22.8	163	17.6	88	15.2	75	21.6
>90	1099	26.8	647	32.1	452	21.7	279	30.1	209	36.1	70	20.2
Non-occupational physical activity, mean (SD)	73.0 (61.0)		81.0 (68.6)		65.3 (51.6)		77.1 (66.1)		83.7 (67.8)		66.1 (61.6)	
Other risk factors												
Ethanol intake (grams/day) [2]												
Abstainers	977	23.8	303	15.0	674	32.4	176	19.0	65	11.2	111	32.0
<5	1170	28.5	422	20.9	748	36.0	236	25.5	109	18.8	127	36.6
5-<15	932	22.7	542	26.8	390	18.8	208	22.5	140	24.2	68	19.6
15-<30	647	15.8	452	22.4	195	9.4	148	16.0	123	21.2	25	7.2
≥30	373	9.1	300	14.9	73	3.5	158	17.1	142	24.5	16	4.6

Exposure variables and potential confounders	Subcohort members						Cancer of Unknown Primary cases					
	Total (M+F) (n=4099)		Men only (n=2019)		Women only (n=2080)		Total (M+F) (n=926)		Men only (n=579)		Women only (n=347)	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Cigarette smoking status												
Never smokers	1494	36.5	274	13.6	1220	58.7	250	27.0	56	9.7	194	55.9
Ex smokers	1464	35.7	1041	51.6	423	20.3	324	35.0	263	45.4	61	17.6
Current smokers	1141	27.8	704	34.9	437	21.0	352	38.0	260	44.9	92	26.5
Frequency of cigarette smoking (N/day), mean (SD) [3]	15.8 (10.1)		17.4 (10.5)		12.4 (8.2)		17.9 (10.5)		19.2 (10.8)		13.4 (7.8)	
Duration of cigarette smoking (years), mean (SD) [3]	31.9 (12.1)		33.7 (11.6)		28.1 (12.2)		35.5 (11.6)		36.9 (10.9)		30.7 (12.7)	
Level of education (years of education)												
Primary	1177	28.9	497	24.7	680	32.8	262	28.6	132	23.0	130	38.0
Lower vocational	892	21.9	409	20.4	483	23.3	191	20.9	110	19.2	81	23.7
Secondary and medium vocational	1434	35.2	712	35.4	722	34.9	329	35.9	226	39.4	103	30.1
University and higher vocational	577	14.1	391	19.5	186	9.0	134	14.6	106	18.5	28	8.2
Diabetes												
Yes	145	3.5	68	3.4	77	3.7	35	3.8	20	3.5	15	4.3
History of cancer in a first degree relative												
Yes	1866	45.5	893	44.2	973	46.8	441	47.6	274	47.3	167	48.1

Notes

1 Analyses of BMI at age 20 years, and change in BMI since age 20 years were based on 772 CUP cases and 3516 subcohort members with complete data.

2 In consumers only

3 In users only

Table 3 Hazard ratios and 95% confidence intervals for anthropometry for men and women individually, and Cancer of Unknown Primary risk in the Netherlands Cohort Study

	Subcohort members		Cancer of Unknown Primary cases		
	Categorical median	Person time at risk (years)	Cases n	Multivariable adjusted HR	95% CI
Height (cm) at baseline [1]					
Men					
<170	167	4 145	91	1	Reference
170-<175	172	8 031	151	0.90	(0.65- 1.24)
175-<180	176	9 341	152	0.85	(0.61- 1.19)
180-<185	182	6 633	119	0.88	(0.61- 1.27)
≥185	187	3 940	66	0.91	(0.59- 1.41)
p for trend [2]				0.67	
Continuous, 5 centimeter increments		32 090	579	0.98	(0.90- 1.08)
Women					
<160	156	6 096	68	1	Reference
160-<165	162	9 460	77	0.74	(0.51- 1.06)
165-<170	167	12 508	105	0.77	(0.54- 1.09)
170-<175	172	6 322	70	1.03	(0.69- 1.51)
≥175	176	2 646	27	0.99	(0.59- 1.67)
p for trend [2]				0.62	
Continuous, 5 centimeter increments		37 032	347	1.04	(0.94- 1.16)
p for interaction [3]				0.041	
Clothing size					
Men (trouser size) [4 5]					
<50		4 648	86	0.94	(0.68- 1.30)
50-51		6 687	133	1	Reference
52-53		10 155	168	0.82	(0.62- 1.08)
54-55		5 492	84	0.76	(0.54- 1.05)
>56		5 109	108	0.93	(0.68- 1.28)
p for trend [2]				0.46	

	Subcohort members		Cancer of Unknown Primary cases		
	Categorical median	Person time at risk (years)	Cases n	Multivariable adjusted HR	95% CI
Women (skirt size) [6]					
<40		6 901	58	0.91	(0.63- 1.30)
42		9 216	87	1	Reference
44		10 111	96	0.99	(0.72- 1.36)
46-48		9 055	78	0.87	(0.62- 1.22)
>50		1 567	25	1.51	(0.91- 2.52)
p for trend [2]				0.55	

Notes

[1] Multivariable analyses of height were adjusted for age at baseline (years;continuous), sex, alcohol consumption (grams of ethanol per day; continuous), cigarette smoking status (never/ever), frequency (continuous; centered), duration (continuous; centered), weight (kilograms; continuous), and cigarette smoking status (never/ever), duration (continuous; centered) as time-varying covariates.

[2] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

[3] *p* value for interaction between sex and height (continuous), based on cross-product terms in the Cox proportional hazards model and Wald test.

[4] Multivariable analyses of trouser size were adjusted for age at baseline (years;continuous), sex, alcohol consumption (grams of ethanol per day; continuous), cigarette smoking status (never/ever), frequency (continuous; centered), duration (continuous; centered), non-occupational physical activity (minutes/day; categories), and cigarette smoking duration (continuous; centered) as a time-varying covariate.

[5] The proportional hazards assumption was violated for the exposure variable in this analysis.

[6] Multivariable analyses of skirt size were adjusted for age at baseline (years;continuous), sex, alcohol consumption (grams of ethanol per day; continuous), cigarette smoking status (never/ever), frequency (continuous; centered), duration (continuous; centered), non-occupational physical activity (minutes/day; categories), and cigarette smoking duration (continuous; centered) as a time-varying covariate.

Table 4 Hazard ratios and 95% confidence intervals for anthropometry for men and women combined, and Cancer of Unknown Primary risk in the Netherlands Cohort Study

	Subcohort members (n=4099)		Cancer of Unknown Primary cases (n=926)		
	Categorical median	Person time at risk (years)	Cases n	Multivariable adjusted HR	95% CI
BMI (kg/m²) at baseline [1]					
<20	19	2 307	26	0.91	(0.58- 1.42)
20-<25	23	35 407	487	1	Reference
25-<30	27	27 021	352	0.90	(0.77- 1.06)
≥30	32	4 387	61	1.11	(0.81- 1.52)
p for trend [2]				0.77	
p for interaction [3]				0.46	
Continuous, 1 kilogram/meter² increments		69 123	926	0.99	(0.97- 1.02)
BMI (kg/m²) at age 20 years [4] [5]					
<20	19	15 351	176	0.89	(0.71- 1.13)
20-<21.5	21	14 374	197	1	Reference
21.5-<23	22	14 379	183	0.93	(0.74- 1.17)
23-<25	24	10 993	151	0.96	(0.75- 1.23)
≥25	26	14 025	219	1.06	(0.85- 1.32)
p for trend [2]				0.25	
p for interaction [3]				0.63	
Continuous, 1 kilogram/meter² increments		59 581	772	1.03	(0.99- 1.06)
Change in BMI (kg/m²) since age 20 years [6] [7]					
<0	-1	6 684	103	1.24	(0.97- 1.61)
0-<4	2	28 767	388	1	Reference
4-<8	6	19 005	215	0.82	(0.68- 0.99)
≥8	9	14 666	220	1.04	(0.86- 1.26)
p for interaction [3]				0.28	

Notes

[1] Multivariable analyses of BMI at baseline were adjusted for age at baseline (years;continuous), sex, alcohol consumption (grams of ethanol per day; continuous), cigarette smoking status (never/ever), frequency (continuous; centered), duration (continuous; centered), non-occupational physical activity (minutes/day; categories), and cigarette smoking status (never/ever), duration (continuous; centered) as time-varying covariates.

[2] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

[3] *p* value for interaction between sex and BMI exposure variables (categorical), based on cross-product terms in the Cox proportional hazards model and Wald test.

[4] Analyses of BMI at age 20 years, and change in BMI since age 20 years were based on 772 CUP cases and 3516 subcohort members with complete data.

[5] Multivariable analyses of BMI at age 20 years were adjusted for age at baseline (years;continuous), sex, alcohol consumption (grams of ethanol per day; continuous), cigarette smoking status (never/ever), frequency (continuous; centered), duration (continuous; centered), non-occupational physical activity (minutes/day; categories), and cigarette smoking status (never/ever), duration (continuous; centered) as time-varying covariates.

[6] Multivariable analyses of change in BMI since age 20 years were adjusted for age at baseline (years;continuous), sex, alcohol consumption (grams of ethanol per day; continuous), cigarette smoking status (never/ever), frequency (continuous; centered), duration (continuous; centered), non-occupational physical activity (minutes/day; categories), BMI at age 20 years (kg/m^2; continuous), and cigarette smoking status (never/ever), duration (continuous; centered) as time-varying covariates.

[7] The proportional hazards assumption was violated for the exposure variable in this analysis.

Table 5 Hazard ratios and 95% confidence intervals for physical activity and Cancer of Unknown Primary risk in the Netherlands Cohort Study

	Subcohort members (n=4099)		Cancer of Unknown Primary cases (n=926)			
	Categorical median	Person time at risk (years)	Cases	Multivariable adjusted		
			n	HR	95% CI	
Non-occupational physical activity (min/day) [1]						
≤30	19	13 915	199	1	Reference	
>30-60	43	21 938	285	0.91	(0.74-	1.12)
>60-90	74	14 732	163	0.85	(0.67-	1.08)
>90	124	18 537	279	0.97	(0.78-	1.20)
p for trend [2]				0.84		
p for interaction [3]				0.71		
Continuous, 30 minutes/day increments		69 123	926	1.01	(0.97-	1.04)

Notes

[1] Multivariable analyses of non-occupational physical activity were adjusted for age at baseline (years;continuous), sex, alcohol consumption (grams of ethanol per day; continuous), cigarette smoking status (never/ever), frequency (continuous; centered), duration (continuous; centered), BMI at baseline (kg/m^2; continuous), and cigarette smoking status (never/ever), duration (continuous; centered) as time-varying covariates.

[2] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

[3] *p* value for interaction between sex and BMI exposure variables (categorical), based on cross-product terms in the Cox proportional hazards model and Wald test.

To the best of our knowledge, this is the first study to have investigated the association between height and CUP risk, and our stratified analyses indicate no association between the height of men or women and CUP risk. Neither the European cohort study nor that conducted in Australia found any association between BMI at baseline and the development of CUP (10, 21). Although their respective categories differed, each concluded that there was no association between BMI at baseline and CUP risk. Unlike those studies, we not only assessed BMI at baseline but also BMI at age 20 years, finding neither variable to be statistically significantly associated with CUP.

In the European cohort study, an association was found between participants with a higher waist circumference and increased CUP risk (comparing highest versus lowest quartiles) (10). In the NLCS, data was unavailable on waist circumference. However, a previous NLCS study had demonstrated clothing size to be a useful measure within the cohort population when there is no specific data on waist circumference (24). Consequently, we used clothing size as a proxy measure for the distribution of abdominal fat. We found trouser size for men not to be associated with CUP risk. We did, however, observe that women with a skirt size of >50 were at an increased risk of CUP, albeit the association was not statistically significant.

The Australian study observed that moderate or vigorous physical activity and sedentary behaviour were not associated with CUP risk (21). Similarly, in the NLCS we also found no association between physical activity and CUP risk either. Whilst the categories to assess the association between physical activity and CUP risk admittedly differed between both studies, their respective results point towards there being no association.

Strengths and limitations

The strengths of the NLCS are its prospective cohort design, extensive cohort of 120,852 participants, large number of CUP cases, and the wide availability of confounders to adjust for in the analyses. Participants were followed for 20.3 years (only one male participant from the subcohort was lost to follow-up). Data on incident CUP cases was provided by the NCR and includes data from both pathology reports and clinical reports (28). Cancer follow-up through record linkage with these registries was at least 96% complete, and thereby minimizing selection bias (29). Cases were registered by trained registry clerks who had access to the medical files and entered data by applying uniform coding rules. The CUP definition used in the current study may however differ from that used in other countries, as the criteria for defining 'CUP' are

heterogeneous. A potential limitation of the current study is that data on all exposure variables are self-reported, which may have resulted in bias due to misclassification. Recall bias may have occurred as weight at age 20 years was asked at baseline. It should be noted, however, that work elsewhere has concluded that self-reported recall of anthropometric measures in early life is highly correlated with prospectively collected data (30). As such, we expect this to be non-differential between CUP cases and subcohort members. This study has only investigated non-occupational physical activity as an indicator for exercise behaviour and any relationship to CUP. Occupational and non-occupational physical activity might be inversely associated and, as a result, occupational physical activity may confound studies of non-occupational exercise (31). The NLCS solely includes data on participants aged 55-69 years at baseline, CUP does however also occur in adolescents and young adults (7). Regrettably, we can therefore not assess the association between anthropometry, physical activity and CUP risk in adolescents and young adults. This could however be a potential direction for future research in a different cohort with a greater range in the age of the participants. Unfortunately, we lack the data to check which diagnostic methods were used for our CUP cases. Nevertheless, if we restrict our analysis to histologically verified CUP cases alone, for whom extended diagnostic methods are more likely, then only minor differences and no overt changes were observed when the results are compared to the overall analyses. Accordingly, the findings from the complete multivariable analyses appear representative for CUP cases with or without an extensive diagnostic work-up.

Conclusions

The findings presented here demonstrate no association between anthropometry, physical activity and the development of CUP. As a result, lifestyle recommendations on cancer prevention regarding maintaining a healthy weight, and being physically active cannot be used in the prevention of CUP.

References

1. Loffler H, Kramer A. Introduction. In: Loffler H, Kramer A, editors. Cancer of Unknown Primary Springer International Publishing Switzerland; 2016.
2. Schroten-Loef C, Verhoeven R, de Hingh I, van de Wouw A, van Laarhoven H, Lemmens V. Unknown primary carcinoma in the Netherlands: decrease in incidence and survival times remain poor between 2000 and 2012. Eur J Cancer. 2018;101:77-86.
3. Pavlidis N, Pentheroudakis G. Cancer of unknown primary site. Lancet. 2012;379(9824):1428-35.
4. Pavlidis N, Khaled H, Gaafar R. A mini review on cancer of unknown primary site: A clinical puzzle for the oncologists. J Adv Res. 2015;6(3):375-82.
5. Comprehensive Cancer Centre the Netherlands. Incidentie Primaire Tumor Onbekend: Integraal kankercentrum Nederland; 2020 [Available from: https://iknl.nl/kankersoorten/primaire-tumor-onbekend/registratie/incidentie].
6. Meijer L, De Peuter R, Van der Zwan JM, Loef C. Primaire Tumor Onbekend: Integraal kankercentrum Nederland; 2020 [Available from: https://www.iknl.nl/getmedia/ea952e4d-9af7-4d59-9f03-865c858bc0cd/Rapport_PrimaireTumorOnbekend_IKNL_2020.pdf].
7. Pavlidis N, Rassy E, Smith-Gagen J. Cancer of unknown primary: Incidence rates, risk factors and survival among adolescents and young adults. Int J Cancer. 2020;146(6):1490-8.
8. Urban D, Rao A, Bressel M, Lawrence Y, Mileshkin L. Cancer of unknown primary: a population-based analysis of temporal change and socioeconomic disparities. Br J Cancer. 2013;109(5):1318-24.
9. Hemminki K, Chen B, Melander O, Manjer J, Hallmans G, Hemminki A. Smoking and body mass index as risk factors for subtypes of cancer of unknown primary. Int J Cancer. 2015;136(1):246-7.
10. Kaaks R, Sookthai D, Hemminki K, Kramer A, Boeing H, Wirfalt E, et al. Risk factors for cancers of unknown primary site: Results from the prospective EPIC cohort. Int J Cancer. 2014;135(10):2475-81.
11. Hemminki K, Försti A, Sundquist K, Li X. Cancer of unknown primary is associated with diabetes. Eur J Cancer Prev. 2016;25(3):246-51.
12. Hermans KEPE, Van den Brandt PA, Loef C, Jansen RLH, Schouten LJ. Alcohol consumption, cigarette smoking, and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study. Submitted, personal communication. 2020.
13. Friedenreich CM, Neilson HK, Lynch BM. State of the epidemiological evidence on physical activity and cancer prevention. Eur J Cancer. 2010;46(14):2593-604.
14. Anderson AS, Key TJ, Norat T, Scoccianti C, Cecchini M, Berrino F, et al. European code against cancer 4th edition: obesity, body fatness and cancer. Cancer Epidemiol 2015;39:S34-S45.
15. Calle EE, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nat Rev Cancer. 2004;4(8):579-91.
16. Hursting SD, DiGiovanni J, Dannenberg AJ, Azrad M, LeRoith D, Demark-Wahnefried W, et al. Obesity, energy balance, and cancer: new opportunities for prevention. 2012;5(11):1260-72.
17. Leitzmann M, Powers H, Anderson AS, Scoccianti C, Berrino F, Boutron-Ruault M-C, et al. European Code against Cancer 4th Edition: Physical activity and cancer. Cancer Epidemiol. 2015;39:S46-S55.

18. World Cancer Research Fund. WCRF Cancer Prevention Recommendations: Word Cancer Research Fund; 2016 [Available from: http://wcrf.org/int/research-we-fund/our-cancer-prevention-recommendations].
19. Shephard RJ, Shek PN. Associations between physical activity and susceptibility to cancer. Sports Med. 1998;26(5):293-315.
20. Nieman DC, Wentz LM. The compelling link between physical activity and the body's defense system. J Sport Health Sci. 2019;8(3):201-17.
21. Vajdic C, Perez-Concha O, Dobbins T, Ward RL, Schaffer AL, van Leeuwen M, et al. Demographic, social and lifestyle risk factors for cancer registry-notified cancer of unknown primary site (CUP). Cancer Epidemiol. 2019;60:156-61.
22. Van den Brandt PA, Goldbohm RA, van 't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol. 1990;43(3):285-95.
23. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol. 1990;19(3):553-8.
24. Hughes LA, Schouten LJ, Goldbohm RA, van den Brandt PA, Weijenberg MP. Self-reported clothing size as a proxy measure for body size. Epidemiology. 2009;20(5):673-6.
25. NEVO-table. Dutch food composition table 1986–1987. Voorlichtingsbureau voor de voeding, The Hague, The Netherlands. 1986.
26. Barlow WE. Robust variance estimation for the case-cohort design. Biometrics. 1994:1064-72.
27. Lin DY, Wei L-J. The robust inference for the Cox proportional hazards model. J Am Stat Assoc. 1989;84(408):1074-8.
28. Van der Sanden G, Coebergh J-W, Schouten L, Visser O, Van Leeuwen F. Cancer incidence in the Netherlands in 1989 and 1990: First results of the nationwide Netherlands Cancer Registry. Eur J Cancer. 1995;31(11):1822-9.
29. Goldbohm RA, Van den Brandt PA, Dorant E. Estimation of the coverage of Dutch municipalities by cancer registries and PALGA based on hospital discharge data. Tijdschr Soc Gezondheidsz 1994;72:80-4.
30. De Rubeis V, Bayat S, Griffith LE, Smith BT, Anderson LN. Validity of self-reported recall of anthropometric measures in early life: A systematic review and meta-analysis. Obes Rev. 2019;20(10):1426-40.
31. Biesma RG, Schouten LJ, Dirx MJ, Goldbohm RA, Van Den Brandt PA. Physical activity and risk of ovarian cancer: results from the Netherlands Cohort Study (The Netherlands). Cancer Causes Control. 2006;17(1):109-15.

CHAPTER 4

VEGETABLE AND FRUIT CONSUMPTION AND CANCER OF UNKNOWN PRIMARY RISK

Karlijn E.P.E. Hermans, Piet A. van den Brandt,
Caroline Loef, Rob L. H. Jansen, and Leo J. Schouten

BMC Cancer. 2022; 22: 399

Abstract

Background: Cancer of Unknown Primary (CUP) is a metastatic cancer for which the primary lesion remains unidentifiable during life and little is also known about the modifiable risk factors that contribute to its development. This study investigates whether vegetables and fruits are associated with CUP risk.

Methods: We used data from the prospective Netherlands Cohort Study on Diet and Cancer which includes 120,852 participants aged between 55-69 years in 1986. All participants completed a self-administered questionnaire on cancer risk factors at baseline. Cancer follow-up was established through record linkage to the Netherlands Cancer Registry and the Dutch Pathology Registry. As a result, 867 incident CUP cases and 4,005 subcohort members were available for case-cohort analyses after 20.3 years of follow-up. Multivariable adjusted hazard ratios were calculated using proportional hazards models.

Results: We observed no associations between total vegetable and fruit consumption (combined or as separate groups) and CUP risk. However, there appeared to be an inverse association between the consumption of raw leafy vegetables and CUP. With respect to individual vegetable and fruit items, we found neither vegetable nor fruit items to be associated with CUP risk.

Conclusions: Overall, vegetable and fruit intake were not associated with CUP incidence within this cohort.

Background

Cancer of Unknown Primary (CUP) is a metastasised malignancy for which the primary tumor origin remains unidentifiable during life (1, 2). A historical study has estimated that CUP accounts for 3-5% of all epithelial tumours (3, 4). In a more recent study, it was observed that CUP incidence has decreased over the last 10-20 years. This decline in CUP incidence was investigated by comparing population-based incidence-rates, and its authors concluded that the decrease could possibly be explained due to advanced imaging and molecular profiling (5). In the Netherlands, the disease accounted for approximately 1,300 incident cases, which represented almost 2% of all new cancer diagnoses in 2018 (6, 7). The median survival of CUP patients is 1.7 months (2000-2012) (2). To prevent CUP, it may be beneficial to identify modifiable lifestyle risk factors that have been associated with other cancers. To date, modifiable risk factors that have been associated with CUP are cigarette smoking, and alcohol consumption (dose-response) (8-11). However, the relationship between diet and CUP has been less studied, especially with respect to plant-based nutrition such as vegetables and fruits.

The World Cancer Research Fund reports that the consumption of vegetables and fruits may reduce cancer risk, although the association may be restricted to specific cancers (12-14). In addition, they describe that non-starchy vegetables and fruits have been linked to protecting against a number of aerodigestive cancers (12, 13). Associations between diet and cancer are complex as each bioactive food constituent has the potential to modify aspects of carcinogenesis, either individually or in combination with several micronutrients (alongside quantity, timing, and duration of exposure to those constituents) (12). Then again, a lower intake of vegetables and fruits (low intake levels of carotenoids, vitamin A, C, E) has been linked to increase levels of oxidative stress and inflammation, alongside genomic instability, reduced apoptosis and increased proliferation (14).

To the best of our knowledge, only one Australian prospective cohort study has investigated the relationship between diet and CUP, in which they did not find any associations between vegetable or fruit consumption and CUP risk (10). However, it should be noted that the study only examined vegetable and fruit consumption by using the usual number of servings as ≥ 5 vegetables/day and ≥ 2 fruits/day in relation to CUP. Similarly, it did not investigate specific groups of vegetables and fruits, nor individual vegetable and fruit items. For that reason, we decided to

investigate the relationship between vegetable and fruit consumption and CUP risk in greater detail by using combined groups of vegetables and fruits, as well as individual vegetable and fruit items. In addition, we aimed to examine residual confounding by cigarette smoking status on the association between vegetable and fruit consumption and CUP risk, as cigarette smoking has been linked to increase CUP risk.

Methods

Study design and population

The prospective Netherlands Cohort Study on Diet and Cancer (NLCS) was started in September 1986 and included 58,279 men and 62,573 women aged between 55-69 years. Participants originated from 204 Dutch computerized municipal population registries. Data processing and analysis were based on the case-cohort design for efficiency reasons. Incident cancer cases were derived from the full cohort while the number of person-years at risk was estimated from a subcohort of 5,000 participants who were randomly sampled from the full cohort immediately after baseline (15). The subcohort comprises a group of participants in which CUP cases can occur (16). The case-cohort design implies that cases can arise both inside and outside the subcohort. The cases in the subcohort are at risk from baseline until cancer incidence, cases outside the subcohort have been assigned a minimal person-time at risk in order to be included in the statistical analysis. Participants who had reported a history of cancer (except for skin cancer) at baseline were excluded from analyses (see Figure 1).

Outcome measure

CUP is defined here as a metastasised epithelial malignancy with no identifiable primary tumor origin after cytological and/or histological verification during a patient's lifetime. This CUP definition only includes epithelial malignancies (ICD-O-3: M-8000 - M-8570) and thus excludes non-epithelial cancers, such as sarcoma, lymphoma, mesothelioma, and melanoma.

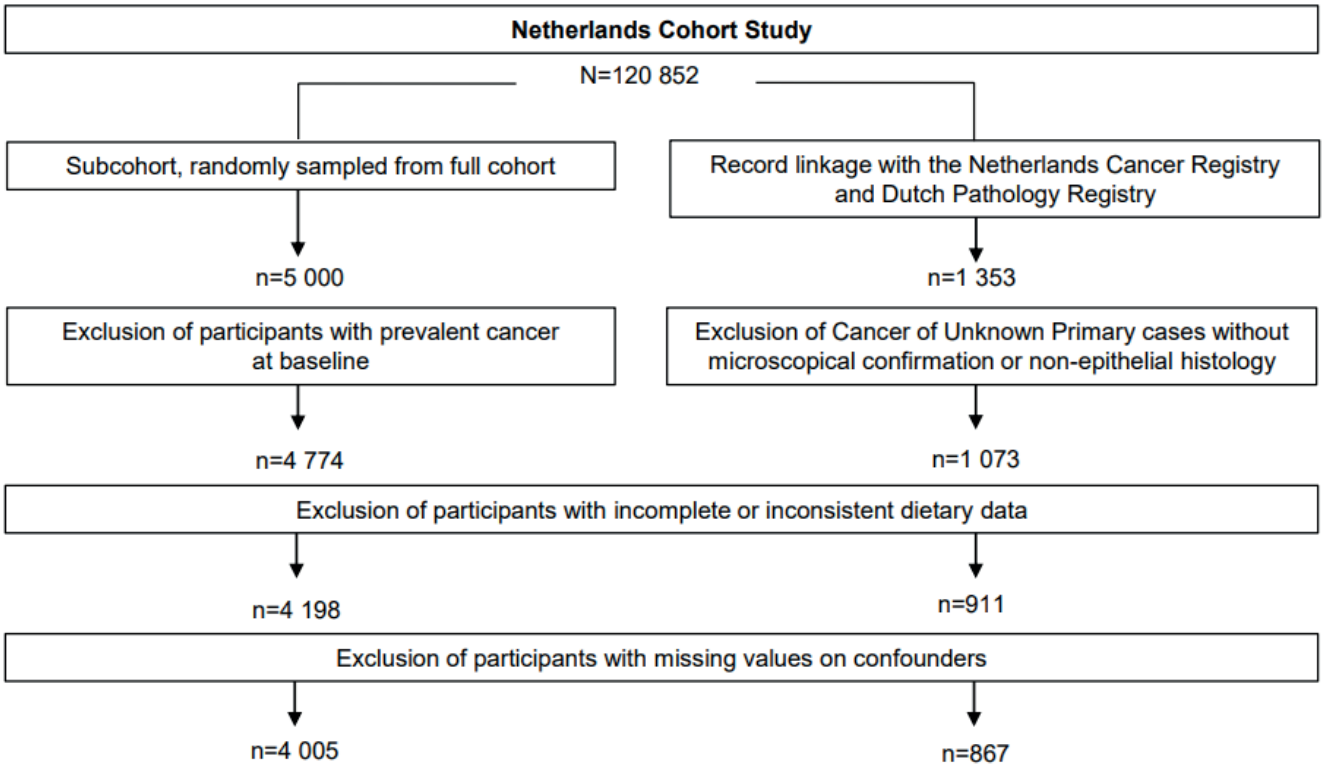


Figure 1 Flow diagram of subcohort members and Cancer of Unknown Primary cases in the Netherlands Cohort Study on whom analyses are based

Follow-up

Cancer follow-up was established through annual record linkage with the Netherlands Cancer Registry (NCR) and the Dutch Pathology Registry (PALGA) (17). Information regarding the site of metastasis was obtained from the NCR, but this data was only partially available and, therefore, supplementary information was retrieved from the pathology excerpts provided by PALGA. These pathology excerpts were also used to determine whether cytological and/or histological confirmed cases had been correctly categorised in the data received from the NCR.

Questionnaire data

All cohort members completed a self-administered questionnaire, which included detailed questions on dietary habits, lifestyle, and other cancer risk factors. The dietary section was a validated 150-item semi quantitative food-frequency questionnaire (FFQ) that concentrated on the habitual consumption of foods and beverages during the year preceding baseline (18). The Spearman correlation coefficient was 0.38 for total vegetable consumption and 0.60 for total fruit consumption, compared to the results of the 9 recording days. The relatively low correlation for total vegetable consumption may derive from lack of variation in consumption and possibly due to imprecise estimation of the portion size (18, 19). Participants were asked to indicate how often they consumed vegetables (15 cooked vegetables, 4 raw vegetables), both in summer and in winter. They were able to choose from one out of six categories: never or less than once a month, 1 time per month, 2 to 3 times per month, 1 time per week, 2 times per week, or 3 to 7 times per week. Usual serving sizes were asked for string beans and cooked endive only; the mean of these values served as an indicator for serving sizes of all cooked vegetables. Participants who did not report their usual serving sizes were assigned a default value. If participants reported only one serving size, then the individual serving size was derived using a conversion factor. Both the default value and the conversion factor were derived from a pilot study (20). Tomato and sweet pepper consumption were asked to be reported in frequency per week and per month, respectively, both in summer and in winter. Participants were asked to indicate how often they consumed fruit by choosing from one out of seven categories: never or less than once a month, 1 time per month, 2 to 3 times per month, 1 time per week, 2 to 3 times per week, 4 to 5 times per week, or 6 to 7 times per week. For all the fruits of interest, participants were able to indicate the amount of each fruit that was

consumed. Frequencies and amounts were converted to grams per day. For both vegetable and fruit consumption, dietary data measured in summer and winter were merged and averaged into specific intake variables for analyses purposes. The questionnaire was also used to measure exposure to tobacco smoking. Tobacco smoking was addressed through questions on baseline smoking status, and the ages at first exposure and last (if stopped) exposure to smoking. Questions were also asked about smoking frequency and smoking duration (excluding stopping periods), for cigarette, cigar, and pipe smokers. Participants who indicated that they had never smoked cigarettes were considered never smokers.

Statistical methods

Person-years at risk were calculated from baseline (17 September 1986) until CUP diagnosis, death, emigration, loss to follow-up, or end of follow-up (31 December 2006), whichever occurred first. Patient characteristics were presented for CUP cases and stratified for histological and cytological confirmation. General characteristics were presented for subcohort members and CUP cases with frequencies (percentages) for categorical variables, and means including standard deviations for continuous variables. Based on the distribution of the subcohort, participants were compared using quartiles (Q) of vegetable, legume, and fruit consumption. For continuous analyses, increments of 25 grams per day were used. The composition of the vegetable, legume, and fruit groups that were studied within the NLCS are described in Table 1.

Vegetable and fruit consumption were mutually adjusted in the analyses, which means that vegetable consumption was additionally adjusted for fruit consumption, whereas fruit consumption was additionally adjusted for vegetable consumption. Legume consumption was additionally adjusted for vegetable and fruit intake. The predefined confounders included: age at baseline (years, continuous); sex (male/female); current cigarette smoking status (never/ever); cigarette smoking frequency (number of cigarettes smoked per day); and cigarette smoking duration (number of years smoking). We included the smoking variables as predefined confounders, as they have been linked to increased CUP risk (8-11). Additionally, smokers have been observed to consume lower amounts of vegetables and fruits in comparison to non-smokers (21). The potential confounders included: alcohol consumption (ethanol intake per day); body mass index (BMI) at baseline (kg/m^2);

non-occupational physical activity (<30 min/day, 30-60 min/day, 60-90 min/day and >90 min/day); socio-economic status (highest level of education); diabetes (yes/no); and history of cancer in a first-degree relative (yes/no). Variables were considered a confounder if they changed the HR by >10%. Accordingly, none of the potential confounders were included in the final model.

Cox proportional hazards models were used to estimate age- and sex-adjusted, and multivariable adjust hazard ratios (HRs) with 95% confidence intervals (CIs). Time since baseline (1986) was used for the time axis. Standard errors were calculated using the robust Huber-White sandwich estimator to account for additional variance introduced by sampling from the full cohort (22). The proportional hazards assumption was tested using the scaled Schoenfeld residuals (23). In cases where the assumption had been violated, a time-varying coefficient for that variable was added to the model where appropriate. Ordinal exposure variables were fitted as continuous variables in trend analyses. Wald tests and cross-product terms were used to evaluate potential multiplicative interaction between total vegetable and fruit consumption (combined and individually), with sex, and CUP risk, and between total vegetable and fruit consumption (combined and individually), cigarette smoking frequency, and CUP risk. Analyses were conducted using Stata version 15. *P* values were considered statistically significant if $p < 0.05$.

We performed three sensitivity analyses. The first sensitivity analysis was restricted to histologically verified CUP cases alone. For this analysis, patients who received a cytological verification alone were excluded. Patients who were histologically verified are more likely to have undergone extensive diagnostic investigation(s) to rule out the primary tumour origin. For those patients who received cytological verification alone, other factors may have played a role in the decision to refrain from further diagnostic investigation, such as age, comorbidities, performance status, localisation of the metastasis, and the patient's decision. The second sensitivity analysis was performed after the first two years of follow-up had been excluded so as to check for potential reverse causality bias as a result of preclinical cancer at baseline. To assess whether associations differed over time, we conducted a third analysis in which we compared the first ten years of follow-up (<1996) to the last ten years of follow-up (≥1996).

Table 1 Composition of vegetable and fruit groups, based on vegetable and fruit items that were asked in the food-frequency questionnaire in the Netherlands Cohort Study

Food group	Composition
Total vegetables	Cooked vegetables plus raw vegetables
Cooked vegetables	Beetroot, broad beans, Brussels sprouts, cauliflower, cabbage (white/green), cooked carrots, cooked endive, kale, leek, mushrooms, onions, rhubarb, sauerkraut, spinach, string beans, sweet peppers and other cooked vegetables originating from an open-ended question on frequently consumed items not listed in the questionnaire
Raw vegetables	Gherkins, lettuce, raw carrots, raw endive, tomatoes and other raw vegetables from an open-ended question on frequently consumed items not listed in the questionnaire
Brassica vegetables	Brussels sprouts, cabbage (white/green), cauliflower and kale
Leafy vegetables, cooked	Cooked endive and spinach
Leafy vegetables, raw	Lettuce and raw endive
Legumes	Broad beans, dried pulses and string beans
Allium vegetables	Leek and onions
Total fruits	Apples/pears, bananas, grapefruits and fresh grapefruit juice, grapes, mandarins, oranges and fresh orange juice, raisins/other dried fruit, strawberries and other fruits originating from an open-ended question on frequently consumed items not listed in the questionnaire
Citrus fruits	Fresh lemon juice, grapefruits and fresh grapefruit juice, mandarins, oranges and fresh orange juice

Results

After 20.3 years of follow-up (17 September 1986 until 31 December 2006), data was available for a total of 1,353 potential CUP cases and 4,774 participants of the subcohort. After excluding CUP cases with neither microscopical confirmation or non-epithelial histology, a total of 1,073 CUP cases remained. Participants with incomplete or inconsistent dietary data were excluded from analyses. This resulted in 867 available CUP cases and 4,005 subcohort members with complete and consistent dietary data. In general, when comparing differences between CUP cases and subcohort members, we observed that CUP cases consumed lower amounts of vegetables (mean values 185.8 versus 189.0 grams per day, respectively) (see Table 2). Male CUP cases in particular consumed lower amounts of vegetables (mean values 182.3 versus 187.0 grams per day, respectively), while female CUP cases consumed a more similar number of vegetables (mean values 191.6 versus

190.9 grams per day, respectively). We also observed that CUP cases consumed lower amounts of fruits (mean values 164.7 versus 175.5 grams per day, respectively).

Results from the age- and sex-adjusted analyses were comparable to the results of the multivariable adjusted analyses. Therefore, we only discuss the multivariable adjusted results. We observed no association between total vegetable and fruit consumption (HR for Q4 vs. Q1: 0.98, 95% CI: 0.92-1.05, P_{trend} = 0.63) and CUP risk (see Table 3). In addition, when mutually adjusted, we found no association between total vegetables (HR for Q4 vs. Q1: 0.87, 95% CI: 0.69-1.09, P_{trend} = 0.38) or total fruits (HR for Q4 vs. Q1: 0.94, 95% CI: 0.75-1.17, P_{trend} = 0.56) and CUP risk. Furthermore, we found no associations between the following vegetable groups: cooked vegetables (HR for Q4 vs. Q1: 1.06, 95% CI: 0.82-1.38, P_{trend} = 0.71), raw vegetables (HR for Q4 vs. Q1: 0.96, 95% CI: 0.75-1.22, P_{trend} = 0.94), legumes (HR for Q4 vs. Q1: 1.21, 95% CI: 0.97-1.52, P_{trend} = 0.14), brassica vegetables (HR for Q4 vs. Q1: 1.01, 95% CI: 0.81-1.27, P_{trend} = 0.92), allium vegetables (HR for Q4 vs. Q1: 1.14, 95% CI: 0.91-1.42, P_{trend} = 0.48), cooked leafy vegetables (HR for Q4 vs. Q1: 0.92, 95% CI: 0.74-1.15, P_{trend} = 0.68), or the fruit group: citrus fruits (HR for Q4 vs. Q1: 1.15, 95% CI: 0.93-1.42, P_{trend} = 0.37) and CUP risk. However, we observed a statistically significant trend between the consumption of raw leafy vegetables and a decreased CUP risk (HR for Q4 vs. Q1: 0.82, 95% CI: 0.64-1.03, P_{trend} = 0.03). With respect to individual vegetable and fruit items, which were mutually adjusted, we found no association between the individual vegetable items or the individual fruit items and the development of CUP (see Table 4).

No multiplicative interactions were observed between sex and the association between total vegetable and fruit consumption (combined), vegetable consumption, or fruit consumption, in relation to CUP risk ($P_{interaction}$ = 0.20, 0.17, and 0.46, respectively). However, we did observe multiplicative interactions between vegetables and fruits (combined), and fruit consumption and smoking status in relation to CUP risk ($P_{interaction}$ = 0.03, 0.02, respectively), but not between vegetable consumption and smoking status in relation to CUP risk ($P_{interaction}$ = 0.67). Furthermore, the potential for residual confounding was evaluated based on cigarette smoking status and the relationship between vegetable and fruit consumption and CUP risk (see Table 5). In current smokers, the association of vegetables and fruits with CUP risk was inverse, although not statistically significant (per 25 grams per day increment HR: 0.89, 95% CI: 0.79-1.00, P_{trend}

= 0.06). In never and ex-smokers, vegetable and fruit consumption was not associated with CUP risk. Furthermore, current smokers with the highest fruit intake compared to the lowest fruit intake appeared to have a reduced CUP risk (HR for Q4 vs. Q1: 0.65, 95% CI: 0.43-0.99, although the P_{trend} = 0.16 was not statistically significant).

Results from all three sensitivity analyses, when restricted to histologically verified CUP cases alone (n=614), after excluding the first two years of follow-up, and when comparing the first ten years of follow-up (<1996) to the last ten years of follow-up (≥1996), did not differ substantially from the findings of the overall analyses (see Supplementary Tables 1-6).

Table 2 Characteristics of Cancer of Unknown Primary cases and subcohort members in the Netherlands Cohort Study at baseline

	Subcohort members				Cancer of Unknown Primary cases			
	(n=4005)				(n=867)			
Characteristic	**n**	**(%)**	**mean**	**SD**	**n**	**(%)**	**mean**	**SD**
Age at baseline (years)								
55-59	1550	38.7			265	30.6		
60-64	1389	34.7			340	39.2		
65-69	1066	26.6			262	30.2		
Sex								
Men	1941	48.5			537	61.9		
Women	2064	51.5			330	38.1		
Total vegetable and fruit consumption (g/day)			364.5	152.4			350.5	145.5
Men			342.4	149.6			329.5	142.4
Women			385.2	152.1			384.7	144.3
Total vegetable consumption (g/day)			189	75.5			185.8	74.2
Men			187	76.0			182.3	75.1
Women			190.9	75.1			191.6	72.5
Total fruit consumption (g/day)			175.5	118.2			164.7	113.8
Men			155.4	114.5			147.3	110.4
Women			194.4	118.5			193.1	113.7
Ethanol intake (grams/day) [1]								
Abstainers	920	23.6			155	18.2		
<5	1105	28.4			220	25.9		
5-<15	896	23.0			196	23.0		
15-<30	623	16.0			136	16.0		
≥30	354	9.1			144	16.9		

Characteristic	Subcohort members (n=4005) n	(%)	mean	SD	Cancer of Unknown Primary cases (n=867) n	(%)	mean	SD
Cigarette smoking status								
Never smokers	1500	37.5			252	29.1		
Ex smokers	1439	35.9			304	35.1		
Current smokers	1066	26.6			311	35.9		
Frequency of cigarette smoking (N/day) [1]			15.7	10.0			17.8	10.1
Duration of cigarette smoking (years) [1]			31.8	12.1			35.3	11.7
Body Mass Index at baseline (kg/m²)			25.0	3.1			24.9	3.0
Non-occupational physical activity (min/day)								
≤30	838	21.2			181	21.2		
>30-60	1240	31.4			261	30.6		
>60-90	834	21.1			154	18.1		
>90	1043	26.4			257	30.1		
Level of education (years of education)								
Primary	1137	28.9			229	26.6		
Lower vocational	857	21.5			172	20.0		
Secondary and medium vocational	1423	35.7			328	38.1		
University and higher vocational	566	14.2			131	15.2		
Diabetes								
Yes	138	3.5			31	3.6		
First grade family history of cancer [3]								
Yes	1836	45.8			422	48.7		

Notes

[1] In consumers only

[2] In users only

[3] First degree relative with cancer

Table 3 Hazard ratios and 95% confidence intervals for vegetable and fruit consumption and Cancer of Unknown Primary risk in the Netherlands Cohort Study

	Subcohort members (n=4005)			Cancer of Unknown Primary cases (n=867)					
	Categorical median (grams per day)		Person time at risk (years)	Cases	Age- and sex - adjusted		Multivariable adjusted [1]		
	Men	Women		n	HR	95% CI	HR	95% CI	
Total vegetables and fruits									
Q1	188	226	16 680	224	1	Reference	1	Reference	
Q2	282	323	16 957	224	0.96	(0.78- 1.19)	1.02	(0.83-	1.27)
Q3	363	411	16 989	209	0.90	(0.72- 1.11)	0.96	(0.78-	1.19)
Q4	496	552	17 184	210	0.87	(0.70- 1.07)	0.97	(0.78-	1.20)
p for trend [2]					0.14		0.63		
Continuous, 25 grams per day increments			67 810	867	0.95	(0.89- 1.02)	0.98	(0.92-	1.05)
Total vegetables [3]									
Q1	109	124	16 600	228	1	Reference	1	Reference	
Q2	156	160	17 022	211	0.91	(0.74- 1.13)	0.94	(0.76-	1.17)
Q3	199	202	17 172	233	0.99	(0.80- 1.22)	1.04	(0.84-	1.28)
Q4	271	277	17 016	195	0.84	(0.68- 1.04)	0.87	(0.69-	1.09)
p for trend [2]					0.21		0.38		
Continuous, 25 grams per day increments			67 810	867	0.96	(0.90- 1.02)	0.97	(0.90-	1.04)
Cooked vegetables [4]									
Q1	85	86	16 707	223	1	Reference	1	Reference	
Q2	125	124	16 976	216	0.96	(0.77- 1.18)	1.00	(0.80-	1.24)
Q3	160	159	17 320	213	0.93	(0.75- 1.15)	0.99	(0.79-	1.24)
Q4	220	216	16 806	215	0.96	(0.78- 1.19)	1.06	(0.82-	1.38)
p for trend [2]					0.69		0.71		
Continuous, 25 grams per day increments			67 810	867	0.99	(0.92- 1.06)	1.02	(0.94-	1.10)
Raw vegetables [4]									
Q1	8	11	16 680	221	1	Reference	1	Reference	
Q2	24	29	16 982	217	0.95	(0.77- 1.17)	1.04	(0.84-	1.29)
Q3	39	45	17 014	235	1.02	(0.83- 1.25)	1.12	(0.90-	1.39)
Q4	67	72	17 134	194	0.85	(0.68- 1.05)	0.96	(0.75-	1.22)
p for trend [2]					0.23		0.94		
Continuous, 25 grams per day increments			67 810	867	0.96	(0.90- 1.03)	0.99	(0.93-	1.07)

	Subcohort members (n=4005)			Cancer of Unknown Primary cases (n=867)					
	Categorical median (grams per day)		Person time at risk (years)	Cases	Age- and sex - adjusted			Multivariable adjusted [1]	
	Men	Women		n	HR	95% CI		HR	95% CI
Legumes [4]									
Q1	13	11	16 934	203	1	Reference		1	Reference
Q2	24	21	17 036	217	1.09	(0.88-	1.35)	1.11	(0.90- 1.38)
Q3	36	32	17 055	214	1.07	(0.86-	1.32)	1.08	(0.87- 1.35)
Q4	57	52	16 784	233	1.20	(0.97-	1.48)	1.21	(0.97- 1.52)
p for trend [2]					0.13			0.14	
Continuous, 25 grams per day increments			67 810	867	1.05	(0.98-	1.13)	1.06	(0.98- 1.14)
Brassica vegetables [4]									
Q1	12	12	16 718	228	1	Reference		1	Reference
Q2	24	23	17 043	214	0.94	(0.76-	1.16)	0.95	(0.77- 1.18)
Q3	35	33	17 162	205	0.88	(0.71-	1.09)	0.89	(0.71- 1.11)
Q4	54	53	16 888	220	0.97	(0.79-	1.20)	1.01	(0.81- 1.27)
p for trend [2]					0.68			0.92	
Continuous, 25 grams per day increments			67 810	867	0.99	(0.92-	1.06)	0.99	(0.92- 1.07)
Allium vegetables [4]									
Q1	6	4	18 455	240	1	Reference		1	Reference
Q2	19	20	15 155	199	1.04	(0.84-	1.28)	1.07	(0.86- 1.33)
Q3	31	33	17 224	195	0.91	(0.74-	1.12)	0.94	(0.75- 1.16)
Q4	55	55	16 975	233	1.06	(0.87-	1.31)	1.14	(0.91- 1.42)
p for trend [2]					0.84			0.48	
Continuous, 25 grams per day increments			67 810	867	1.01	(0.94-	1.08)	1.03	(0.96- 1.10)
Leafy vegetables, cooked [4]									
Q1	5	5	16 925	232	1	Reference		1	Reference
Q2	15	15	16 985	211	0.89	(0.72-	1.09)	0.89	(0.72- 1.11)
Q3	24	24	17 051	218	0.95	(0.77-	1.17)	0.99	(0.80- 1.22)
Q4	39	38	16 849	206	0.90	(0.72-	1.10)	0.92	(0.74- 1.15)
p for trend [2]					0.40			0.68	
Continuous, 25 grams per day increments			67 810	867	0.97	(0.91-	1.04)	0.99	(0.92- 1.06)

	Subcohort members (n=4005)			Cancer of Unknown Primary cases (n=867)					
	Categorical median (grams per day)		Person time at risk (years)	Cases	Age- and sex - adjusted			Multivariable adjusted [1]	
	Men	Women		n	HR	95% CI		HR	95% CI
Leafy vegetables, raw [4]									
Q1	1	1	12 911	197	1	Reference		1	Reference
Q2	4	4	15 347	217	0.96	(0.77-	1.20)	0.98	(0.78- 1.22)
Q3	9	9	21 890	252	0.77	(0.62-	0.95)	0.80	(0.64- 0.99)
Q4	20	20	17 661	201	0.77	(0.61-	0.96)	0.82	(0.64- 1.03)
p for trend [2]					0.004			0.03	
Continuous, 25 grams per day increments			67 810	867	0.90	(0.84-	0.97)	0.92	(0.85- 0.99)
Total fruits [5]									
Q1	41	74	16 675	236	1	Reference		1	Reference
Q2	109	144	16 980	216	0.88	(0.71-	1.09)	0.94	(0.76- 1.16)
Q3	165	210	17 040	205	0.81	(0.66-	1.00)	0.92	(0.74- 1.15)
Q4	270	326	17 115	210	0.82	(0.66-	1.01)	0.94	(0.75- 1.17)
p for trend [2]					0.05			0.56	
Continuous, 25 grams per day increments			67 810	867	0.93	(0.87-	0.99)	0.98	(0.91- 1.05)
Citrus fruits [5]									
Q1	0	6	16 947	222	1	Reference		1	Reference
Q2	15	36	17 118	213	0.93	(0.75-	1.15)	0.98	(0.79- 1.21)
Q3	52	83	16 845	180	0.77	(0.62-	0.96)	0.85	(0.68- 1.06)
Q4	115	167	16 900	252	1.07	(0.87-	1.31)	1.15	(0.93- 1.42)
p for trend [2]					0.84			0.37	
Continuous, 25 grams per day increments			67 810	867	1.01	(0.94-	1.08)	1.03	(0.96- 1.11)

Notes

[1] Analyses were adjusted for age at baseline (years), sex, cigarette smoking status (never/ever), frequency (continuous; centered), and duration (continuous; centered). Additionally adjusted for cigarette smoking status (never/ever), and duration (continuous; centered) as time-varying covariates.

[2] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

[3] Additionally adjusted for total fruit consumption (grams per day; continuous).

[4] Additionally adjusted for total vegetable and fruit consumption (grams per day; continuous).

[5] Additionally adjusted for total vegetable consumption (grams per day; continuous).

Table 4 Hazard ratios and 95% confidence intervals for individual vegetable and fruit items and Cancer of Unknown Primary risk in the Netherlands Cohort Study [1]

		Cancer of Unknown Primary cases (n=867)				
		Follow-up time (years)	Age- and sex- adjusted		Multivariable adjusted [2]	
			HR	95% CI	HR	95% CI
Vegetable item (25 grams per day increments)						
	String/French beans	20.3	1.02	1.15)	1.01	(0.89- 1.15)
	Cauliflower	20.3	0.95	1.14)	0.95	(0.79- 1.15)
	Lettuce	20.3	0.75	1.01)	0.83	(0.62- 1.13)
	Carrots, cooked [3]	0-10	0.95	1.31)	1.03	(0.75- 1.41)
		10-20.3	0.73	0.97)	0.78	(0.59- 1.03)
	Endive, cooked [3]	0-10	0.99	1.31)	1.01	(0.76- 1.33)
		10-20.3	0.83	1.06)	0.85	(0.67- 1.08)
	Brussels sprouts	20.3	1.04	1.35)	1.06	(0.81- 1.37)
	Sauerkraut	20.3	1.07	1.52)	1.12	(0.78- 1.62)
	Tomatoes	20.3	0.96	1.06)	0.98	(0.89- 1.08)
	Onion	20.3	0.99	1.10)	1.02	(0.91- 1.13)
	Spinach	20.3	0.99	1.22)	1.02	(0.82- 1.27)
	Beetroot [3]	0-10	0.91	1.28)	0.99	(0.69- 1.41)
		10-20.3	0.60	0.85)	0.64	(0.44- 0.92)
	Kale	20.3	0.86	1.52)	0.93	(0.53- 1.63)
Fruit item (25 grams per day increments)						
	Apples and pears [3]	0-10	0.95	0.99)	0.97	(0.94- 1.01)
		10-20.3	0.98	1.01)	0.99	(0.96- 1.03)
	Strawberries	20.3	0.99	1.27)	1.06	(0.83- 1.36)
	Oranges and fresh orange juice	20.3	1.01	1.04)	1.03	(0.99- 1.07)

Notes

1 The total person time at risk in the subcohort was 67,810 years.

2 Analyses were adjusted for age at baseline (years), sex, cigarette smoking status (never/ever), frequency (continuous; centered), duration (continuous; centered), and total vegetable and fruit consumption (grams per day; continuous). All items were assessed while additionally using cigarette smoking status (never/ever), and duration (continuous; centered) as time-varying covariates.

3 The proportional hazards assumption was violated for the exposure variable in this analysis, consequently these associations were splitted based on follow-up time.

Table 5 Hazard ratios and 95% confidence intervals for vegetable and fruit consumption and Cancer of Unknown Primary risk in the Netherlands Cohort Study, stratified for cigarette smoking status

	Never smokers				Ex smokers				Current smokers				
	Subcohort members (n=1500)	Cancer of Unknown Primary cases (n=252)			Subcohort members (n=1439)	Cancer of Unknown Primary cases (n=304)			Subcohort members (n=1066)	Cancer of Unknown Primary cases (n=311)			
	Person time at risk (years)	#	Age- and sex-adjusted [1]		Person time at risk (years)	#	Age- and sex-adjusted [1]		Person time at risk (years)	#	Age- and sex-adjusted [1]		*p* for interaction [2]
		n	HR	95% CI		n	HR	95% CI		n	HR	95% CI	
Total vegetables and fruits													0.032
Q1	6 185	53	1	Reference	5 006	59	1	Reference	5 489	112	1	Reference	
Q2	6 470	66	1.17	(0.79- 1.74)	6 227	76	1.05	(0.71- 1.54)	4 260	82	0.95	(0.68- 1.34)	
Q3	6 846	70	1.19	(0.81- 1.76)	6 201	65	0.95	(0.64- 1.40)	3 941	74	0.86	(0.61- 1.22)	
Q4	7 435	63	0.92	(0.62- 1.38)	6 677	104	1.35	(0.94- 1.94)	3 072	43	0.67	(0.45- 1.01)	
p for trend [3]			0.66				0.13				0.06		
Continuous, 25 grams per day increments	26 935	252	0.97	(0.87- 1.09)	24 112	304	1.10	(0.97- 1.24)	16 763	311	0.89	(0.79- 1.00)	
Total vegetables													0.673
Q1	7 081	63	1	Reference	5 102	69	1	Reference	4 417	96	1	Reference	
Q2	6 800	64	1.10	(0.75- 1.60)	6 049	72	0.88	(0.61- 1.29)	4 172	75	0.92	(0.64- 1.32)	
Q3	6 751	71	1.25	(0.86- 1.81)	6 294	88	1.08	(0.75- 1.55)	4 127	74	0.88	(0.61- 1.26)	
Q4	6 303	54	1.04	(0.69- 1.56)	6 666	75	0.79	(0.54- 1.16)	4 046	66	0.81	(0.55- 1.19)	
p for trend [3]			0.65				0.38				0.27		
Continuous, 25 grams per day increments	26 935	252	1.03	(0.91- 1.16)	24 112	304	0.95	(0.85- 1.07)	16 763	311	0.93	(0.83- 1.05)	
Total fruits													0.019
Q1	5 650	53	1	Reference	5 144	67	1	Reference	5 881	116	1	Reference	
Q2	6 415	68	1.15	(0.77- 1.70)	5 980	65	0.83	(0.56- 1.22)	4 585	83	0.91	(0.65- 1.27)	
Q3	7 150	59	0.84	(0.56- 1.27)	6 616	74	0.91	(0.62- 1.32)	3 274	72	1.10	(0.76- 1.56)	
Q4	7 720	72	0.93	(0.62- 1.38)	6 372	98	1.22	(0.84- 1.77)	3 023	40	0.65	(0.43- 0.99)	
p for trend [3]			0.39				0.20				0.16		
Continuous, 25 grams per day increments	26 935	252	0.95	(0.83- 1.07)	24 112	304	1.09	(0.96- 1.23)	16 763	311	0.92	(0.82- 1.03)	

Notes

[1] Analyses were adjusted for age at baseline (years) and sex.

[2] Interactions were calculated with respect to smoking status in relation to the vegetable/fruit variable of interest and CUP risk.

[3] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

Discussion

We have presented here a detailed investigation of the relationship between vegetable and fruit consumption and the development of CUP, which we accomplished by assessing combined groups of vegetables and fruits as well as individual vegetable and fruit items. Our results demonstrate that consuming vegetables and fruits is generally unrelated to CUP incidence within this cohort; however, the consumption of raw leafy vegetables did appear to be associated with a decreased CUP risk. We found no multiplicative interaction between sex in relation to the association between total vegetable and fruit consumption and CUP risk. Yet, we did observe multiplicative interactions between total vegetables and fruits (combined), and fruit consumption and smoking status in relation to CUP risk, but not between vegetable consumption and smoking status in relation to CUP risk.

The Australian cohort study, mentioned in the introduction, investigated the relationship between consuming vegetables and fruits and the risk of developing CUP by comparing 327 incident CUP cases to two randomly selected sets of controls (3:1) using incidence density sampling with replacement (10). It found no relation by assessing plant-based food consumption and the usual number of servings as ≥ 5 vegetables/day and ≥ 2 fruits/day, compared to consuming < 5 vegetables/day and < 2 fruits/day (10). Although the categories differ between the Australian study and those of the NLCS, the respective findings are comparable. Moreover, having analysed combined groups of vegetables and fruits as well as individual vegetable and fruit items in greater detail, we conclude that there is no association between vegetable and fruit consumption and CUP risk. We did, however, observe an inverse association between the consumption of raw leafy vegetables and CUP risk, but this might be a chance finding due to multiple comparisons. As described elsewhere, vegetable and fruit consumption have been associated with a protective effect against cancer, but the association may be restricted to specific cancers (12). Nonetheless, it should be acknowledged that CUP constitutes a group of heterogeneous metastatic cancers, therefore, specific effects from vegetables and/or fruits could be masked.

In an additional analysis, residual confounding by cigarette smoking status was evaluated for its possible influence on the association between vegetable and fruit consumption and CUP risk. We observed no associations for never or ex-smokers

who consumed vegetables and fruits in relation to CUP risk, while current smokers appeared to have a decreased CUP risk, although not statistically significant. This effect may derive from residual confounding by smoking. Our finding is in line with the limited-suggestive evidence by the World Cancer Research Fund that describes the consumption of non-starchy vegetables and fruit to be linked to reduced lung cancer risk in people who smoke or used to smoke tobacco (13).

Strengths and limitations

The strengths of this study are its prospective cohort design, its large cohort population including 120,852 participants, its large number of 867 incident CUP cases, and its ability to correct for multiple and detailed confounders in the analyses. Data on incident CUP cases were provided by the NCR and included information from both pathology reports and clinical reports (24). Pathology excerpts were available to confirm whether the cytological and/or histological confirmed cases had been correctly categorised in the data received from the NCR. Cancer follow-up through record linkage with the NCR and PALGA was at least 96% complete, thereby minimizing selection bias (25). Cases were registered by trained NCR registry clerks who had access to the medical files and who entered data by applying uniform coding rules. It should, however, be acknowledged that we utilised a CUP definition that may differ from that used in other countries, as the criteria for defining 'CUP' are heterogeneous. Another possible limitation is that exposure data were only measured once at baseline in 1986. Vegetable and fruit consumption (both in summer and in winter) were, however, extensively addressed in the FFQ, and we expect that participants in the studied age group (55-69) had stable dietary habits at baseline. The reproducibility of the FFQ as well as the stability of dietary habits as estimated by the test-retest r, was on average 0.07 for nutrients over a time period of five years (26). Nonetheless, it is possible that participants subsequently changed their dietary habits. If they did change their habits, that may have resulted in bias due to misclassification and may have led to underestimation of the effect of vegetable and fruit consumption on CUP risk. We do expect this bias to be non-differential between CUP cases and subcohort members. Unfortunately, we do not have data to check which diagnostic methods were used to identify the primary tumour origin. Nevertheless, if we restrict our analysis to histologically verified CUP cases alone, for whom extended diagnostic methods are more likely, we find that the results do not differ greatly from the

overall multivariable analyses. Accordingly, we can assume that the findings from the overall multivariable analyses are representative of CUP cases with or without an extensive diagnostic work-up. We were unable to conduct subgroup analyses based on histopathological findings as precision medicine was not yet available at the time of the follow-up of our study. Studies with more recent data on CUP cases would therefore be encouraged to conduct such analyses.

Conclusions

In our study, we observed no associations between total vegetable and fruit consumption, total vegetables, cooked vegetables, raw vegetables, legumes, brassica vegetables, allium vegetables, cooked leafy vegetables, total fruits, citrus fruits, and the development of CUP. However, the consumption of raw leafy vegetables appeared to decrease risk of the malignancy. With respect to individual vegetable and fruit items, neither vegetable nor fruit items were found to be associated with CUP risk. We thus conclude that consuming vegetables and fruits is unrelated to CUP incidence within this cohort.

References

1. Loffler H, Kramer A. Introduction. In: Loffler H, Kramer A, editors. Cancer of Unknown Primary Springer International Publishing Switzerland; 2016.
2. Schroten-Loef C, Verhoeven R, de Hingh I, van de Wouw A, van Laarhoven H, Lemmens V. Unknown primary carcinoma in the Netherlands: decrease in incidence and survival times remain poor between 2000 and 2012. Eur J Cancer. 2018;101:77-86.
3. Pavlidis N, Fizazi K. Carcinoma of unknown primary (CUP). Crit Rev Oncol Hematol. 2009;69(3):271-8.
4. Pavlidis N, Khaled H, Gaafar R. A mini review on cancer of unknown primary site: A clinical puzzle for the oncologists. J Adv Res. 2015;6(3):375-82.
5. Rassy E, Pavlidis N. The currently declining incidence of cancer of unknown primary. J Cancer Epidemiol. 2019;61:139-41.
6. Comprehensive Cancer Centre the Netherlands. Incidentie Primaire Tumor Onbekend: Integraal kankercentrum Nederland; 2020 [Available from: https://iknl.nl/kankersoorten/primaire-tumor-onbekend/registratie/incidentie].
7. Meijer L, De Peuter R, Van der Zwan JM, Loef C. Primaire Tumor Onbekend: Integraal kankercentrum Nederland; 2020 [Available from: https://www.iknl.nl/getmedia/ea952e4d-9af7-4d59-9f03-865c858bc0cd/Rapport_PrimaireTumorOnbekend_IKNL_2020.pdf].
8. Hemminki K, Chen B, Melander O, Manjer J, Hallmans G, Hemminki A. Smoking and body mass index as risk factors for subtypes of cancer of unknown primary. Int J Cancer. 2015;136(1):246-7.
9. Kaaks R, Sookthai D, Hemminki K, Kramer A, Boeing H, Wirfalt E, et al. Risk factors for cancers of unknown primary site: Results from the prospective EPIC cohort. Int J Cancer. 2014;135(10):2475-81.
10. Vajdic C, Perez-Concha O, Dobbins T, Ward RL, Schaffer AL, van Leeuwen M, et al. Demographic, social and lifestyle risk factors for cancer registry-notified cancer of unknown primary site (CUP). Cancer Epidemiol. 2019;60:156-61.
11. Hermans KE, van den Brandt PA, Loef C, Jansen RL, Schouten LJ. Alcohol consumption, cigarette smoking and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study. Int J Cancer. 2021.
12. Norat T, Scoccianti C, Boutron-Ruault M-C, Anderson A, Berrino F, Cecchini M, et al. European code against cancer 4th edition: diet and cancer. Cancer Epidemiol. 2015;39:S56-S66.
13. World Cancer Research Fund/American Institute for Cancer Research. Wholegrains, vegetables and fruit and the risk of cancer. 2018.
14. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. 2018.
15. Van den Brandt PA, Goldbohm RA, van 't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol. 1990;43(3):285-95.
16. Barlow WE, Ichikawa L, Rosner D, Izumi SJJoce. Analysis of case-cohort designs. 1999;52(12):1165-72.

17. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol. 1990;19(3):553-8.
18. Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, et al. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin Nutr. 1994;48(4):253-65.
19. Maasland DH, van den Brandt PA, Kremer B, Goldbohm RA, Schouten LJ. Consumption of vegetables and fruits and risk of subtypes of head-neck cancer in the Netherlands Cohort Study. Int J Cancer. 2015;136(5):E396-409.
20. Voorrips LE, Goldbohm RA, Verhoeven DT, van Poppel GA, Sturmans F, Hermus RJ, et al. Vegetable and fruit consumption and lung cancer risk in the Netherlands Cohort Study on diet and cancer. Cancer Causes Control. 2000;11(2):101-15.
21. Birkett NJ. Intake of fruits and vegetables in smokers. Public Health Nutr. 1999;2(2):217-22.
22. Barlow WE. Robust variance estimation for the case-cohort design. Biometrics. 1994:1064-72.
23. Lin DY, Wei L-J. The robust inference for the Cox proportional hazards model. J Am Stat Assoc. 1989;84(408):1074-8.
24. Van der Sanden G, Coebergh J-W, Schouten L, Visser O, Van Leeuwen F. Cancer incidence in the Netherlands in 1989 and 1990: First results of the nationwide Netherlands Cancer Registry. Eur J Cancer. 1995;31(11):1822-9.
25. Goldbohm RA, Van den Brandt PA, Dorant E. Estimation of the coverage of Dutch municipalities by cancer registries and PALGA based on hospital discharge data. Tijdschr Soc Gezondheidsz 1994;72:80-4.
26. Goldbohm RA, van 't Veer P, van den Brandt PA, van 't Hof MA, Brants HA, Sturmans F, et al. Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. Eur J Clin Nutr. 1995;49(6):420-9.

5

CHAPTER 5

MEAT CONSUMPTION AND CANCER OF UNKNOWN PRIMARY RISK

Karlijn E.P.E. Hermans, Piet A. van den Brandt,
Caroline Loef, Rob L. H. Jansen, and Leo J. Schouten

Eur. J. Nutr. 2021

Abstract

Purpose: Cancer of Unknown Primary (CUP) is a metastasised cancer for which no primary lesion could be identified during life. Research into CUP aetiology with respect to dietary factors is particularly scarce. This study investigates whether meat consumption is associated with CUP risk.

Methods: Data was utilised from the prospective Netherlands Cohort Study that includes 120,852 participants aged 55-69 years. All participants completed a self-administered questionnaire on diet and other cancer risk factors at baseline. Cancer follow-up was established through record linkage to the Netherlands Cancer Registry and the Dutch Pathology Registry. A total of 899 CUP cases and 4,111 subcohort members with complete and consistent dietary data were available for case-cohort analyses after 20.3 years of follow-up. Multivariable adjusted hazard ratios (HRs) were calculated using proportional hazards models.

Results: We found a statistically significant positive association with beef and processed meat consumption and CUP risk in women (multivariable adjusted HR Q4 vs. Q1: 1.47, 95% CI: 1.04-2.07, P_{trend} = 0.004 & Q4 vs. Q1: 1.53, 95% CI: 1.08-2.16, P_{trend} = 0.001, respectively), and a non-significant positive association with processed meat consumption and CUP risk in men (multivariable adjusted HR Q4 vs. Q1: 1.33, 95% CI: 0.99-1.79, P_{trend} = 0.15). No associations were observed between red meat (overall), poultry or fish consumption and CUP risk.

Conclusion: In this cohort, beef and processed meat consumption were positively associated with increased CUP risk in women, whereas a non-significant positive association was observed between processed meat consumption and CUP risk in men.

Introduction

Cancer of Unknown Primary (CUP) is a metastasised malignancy for which the primary tumour origin remains unidentifiable during a patient's lifetime [1,2]. It ranks fourth in the most common metastasised cancers in the Netherlands, and with slightly more than 1,300 incident cases in 2018, CUP accounted for almost 2% of all new cancer diagnoses in that year [3,4]. Globally, the median survival for CUP patients is only three months, dependent on available diagnostics as well as incidence and patient characteristics (favourable or unfavourable prognosis, 20-80% respectively) [5-7]. For most CUP patients, curative treatment(s) may no longer be an option [8]. By assessing lifestyle factors that are potentially associated with the disease, however, it may be possible to prevent future CUP patients. Certain modifiable risk factors, such as cigarette smoking and alcohol consumption, have been linked to the development of CUP [9-12]. Nonetheless, the relationship between diet and CUP has been less well studied, and that is particularly true with respect to meat consumption [11].

The consumption of red meat and processed meat has been linked to several types of cancer in previous studies [13]. Indeed, the weight of evidence is such that the International Agency for Research on Cancer (IARC) describes red meats as "probably carcinogenic to humans", and there is also sufficient evidence to classify processed meats as "carcinogenic to humans" [13]. Red meats are unprocessed mammalian muscle meat that contain proteins and important micronutrients such as B vitamins, iron, and zinc [13,14]. Processed meats, by contrast, are those meats that have been transformed through salting, curing, fermentation, smoking, or other processes so as to enhance their flavour or improve their preservation [13]. When those meats are being processed, it can lead to the formation of carcinogenic chemicals (including N-nitroso-compounds (NOC) and polycyclic aromatic hydrocarbons (PAH)) [15,13]. Additionally, the cooking of processed meat (fried, grilled, roasted, boiled and smoked), temperature and duration of cooking, type of fuel used for cooking, and proximity and direct contact with the heat source, can produce known or suspected carcinogens, including heterocyclic aromatic amines (HAA) and PAH [15,13]. While the connection between consuming red meat and processed meat and developing cancer appears rather consistent, the connection between consuming poultry and fish and developing cancer is much less clear. Fish consumption has, however, been linked to anti-inflammatory and anticarcinogenic effects of long-chain *n-3* fatty acids and could thus be beneficial for inhibiting carcinogenesis [16].

The IARC Monographs Working Group has evaluated the consumption of red meat and processed meat with respect to carcinogenicity to humans. Based on epidemiological evidence, it concluded that there are convincing associations between the consumption of red meat and cancer, particularly for cancers of the colorectum, pancreas and prostate [13]. In addition, the consumption of processed meat has been linked to cancers of the colorectum and stomach [13]. The 2018 Continuous Update Project Expert Report of the World Cancer Research Fund (WCRF) and American Institute for Cancer Research (AICR) concluded that the data to study the relation between poultry and cancer risk was "too low quality or too inconsistent, or the number of studies too few, to allow conclusions". For fish consumption, they summarized a 'limited to suggestive' decreased risk of cancers of the colorectum and liver [17].

The relationship between meat consumption and CUP has been investigated in one Australian prospective cohort study [11]. Its authors found no association between red meat consumption and CUP risk, though they did observe a slightly increased risk between processed meat consumption and CUP risk, albeit this was not deemed statistically significant [11]. The current study assesses the association between meat consumption and CUP risk in greater depth by assessing combined groups of meats such as red meat, processed meat, poultry, and fish, as well as individual meat items. Additionally, we investigated whether sex or cigarette smoking status influence the association between meat consumption and CUP risk, by testing multiplicative interactions.

Materials and Methods

Design and study population

The Netherlands Cohort Study on Diet and Cancer (NLCS) includes 120,852 participants aged 55-69 years from 204 Dutch municipalities. The case-cohort design was applied for data processing and analysis. Cases were derived from the full cohort, while the number of person-years at risk for the full cohort was estimated from a subcohort of 5,000 participants who were randomly sampled from the full cohort at baseline in 1986 [18].

Outcome measure

CUP is defined as a metastasised epithelial malignancy with no identifiable primary tumour origin after cytological and/or histological verification during a patient's lifetime. This CUP definition only includes epithelial malignancies (ICD-O-3: M-8000 - M-8570), which excludes for example sarcoma, lymphoma, mesothelioma, and melanoma.

Follow-up

Cancer follow-up was established through annual record linkage with the Netherlands Cancer Registry (NCR) and the Dutch Pathology Registry (PALGA) [19]. Information regarding the site of metastasis was obtained from the NCR, but this data was only partially available and, therefore, supplementary information was retrieved from the pathology excerpts provided by PALGA. These pathology excerpts were also used to determine whether cytological and/or histological confirmed cases had been correctly categorised in the data received from the NCR. After 20.3 years of follow-up (17 September 1986 until 31 December 2006), data was available for a total of 1,353 potential CUP cases, and a subcohort of 4,774 participants after removing members who reported a history of cancer (except for skin cancer) at baseline. After excluding CUP cases without microscopical confirmation or non-epithelial histology, a total of 1,073 CUP cases remained. CUP cases were further subdivided according to histology, according to the number of metastases (multiple metastases of the same type were counted as one metastatic site, for example, bone metastases in hip and vertebra were counted as one), according to localisation of metastasis (up to four locations), and according to survival duration. Participants were removed from the analysis if there was incomplete or inconsistent dietary data, or if there were selected confounders with missing values. As a result, 899 CUP cases and 4,111 subcohort members were available for assessment (see Figure 1).

Questionnaire data

Participants completed a self-administered questionnaire that included detailed questions on dietary habits, lifestyle, and other cancer risk factors. A 150-item semi quantitative food-frequency questionnaire was used that concentrated on the habitual consumption of food and beverages during the year preceding baseline

[20,21]. The food-frequency questionnaire had been validated against a 9-day diet record and was tested for reproducibility in the NLCS [22,23]. The Spearman correlation coefficients for the validity of red meat, processed meat, and fish, as investigated by the questionnaire were, 0.46, 0.54 and 0.53, respectively, compared to the results of the 9-day diet record [22]. The questionnaire contained 14 items on the consumption of meat as the main meal, 5 items on the consumption of meat used as a sandwich filling, and 3 items on the consumption of fish. Meats were grouped into red meat (overall), processed meat, and poultry. Red meat included beef, pork, minced meat (beef and pork), liver, and other meats (e.g., horsemeat, lamb). Processed meat (meat items that had undergone some form of preservation with nitrite salt, fermentation, or smoking) included ham, bacon, smoked beef, pork loin roll, and other sliced cold meats (e.g., sausages). Poultry included chicken and turkey. Fish consumption was measured in relation to the main meal, lunch, or as a snack between meals.

Statistical analysis

Person-years at risk were calculated from baseline (17 September 1986) until CUP diagnosis, death, emigration, loss to follow-up, or end of follow-up (31 December 2006), whichever occurred first. General characteristics were presented for subcohort members and CUP cases with frequencies (percentages) for categorical variables and means including standard deviations for continuous variables. Based on the distribution of the subcohort, participants were compared using quartiles (Q) or categories of red meat, processed meat, poultry, and fish consumption. For continuous analyses, increments of 50 grams per day were used for red meat, beef, pork, minced meat, and poultry consumption, and increments of 25 grams per day were used for liver, processed meat, and fish consumption.

The predefined confounders included: age at baseline (years; continuous); sex (male/female); alcohol consumption (ethanol intake per day); cigarette smoking status (never/ever); cigarette smoking frequency (number of cigarettes smoked per day); cigarette smoking duration (number of years smoking); and total energy intake (kcal/day). The potential confounders included: body mass index (BMI) at baseline (kg/m^2); non-occupational physical activity (<30 min/day, 30-60 min/day, 60-90 min/day and >90 min/day); socio-economic status (highest level of education); diabetes (yes/no); history of cancer in a first degree relative (yes/no);

and vegetable and fruit consumption (grams per day). Variables were considered a confounder if they changed the HR by >10%. Accordingly, none of the potential confounders were included in the final model. No mutual adjustments were conducted between meat groups, as there was insufficient scientific evidence to conclude that they were related to CUP development.

Cox proportional hazards models were utilised to estimate age- and sex-adjusted, and multivariable-adjusted hazard ratios (HRs) with 95% confidence intervals (CIs). Standard errors were calculated using the robust Huber-White sandwich estimator to account for additional variance introduced by sampling from the full cohort [24]. The proportional hazards assumption was tested using the scaled Schoenfeld residuals [25]. In cases where the assumption had been violated, a time-varying covariate for that variable was added to the model where appropriate. Ordinal exposure variables were fitted as continuous variables in trend analyses. Wald tests and cross-product terms were used to evaluate possible multiplicative interaction between sex in relation to meat consumption and CUP risk, or between cigarette smoking status in relation to meat consumption and CUP risk. Analyses were conducted using Stata version 15. *P* values were considered statistically significant if $p < 0.05$.

Three sensitivity analyses were conducted, the first of which was restricted to histologically verified CUP cases only, since it is more likely that those cases had undergone extensive diagnostic investigation(s) to rule out the primary tumour origin. For those patients who received cytological verification alone, other factors may have played a role in the decision to refrain from further diagnostic investigation such as age, comorbidities, performance status, localisation of metastasis, or the patient's decision. The second sensitivity analysis was performed after the first two years of follow-up had been excluded so as to check for potential reverse causality bias as a result of preclinical cancer at baseline. Reverse causality bias may occur if participants change their dietary behaviour as a result of symptoms of preclinical cancer, whereas we are interested to see if dietary behaviour reduces or increases CUP risk. In the third sensitivity analysis, the first ten years of follow-up (<1996) were compared to the last ten years of follow-up (≥1996), as to see whether associations between meat consumption and CUP risk differed over time.

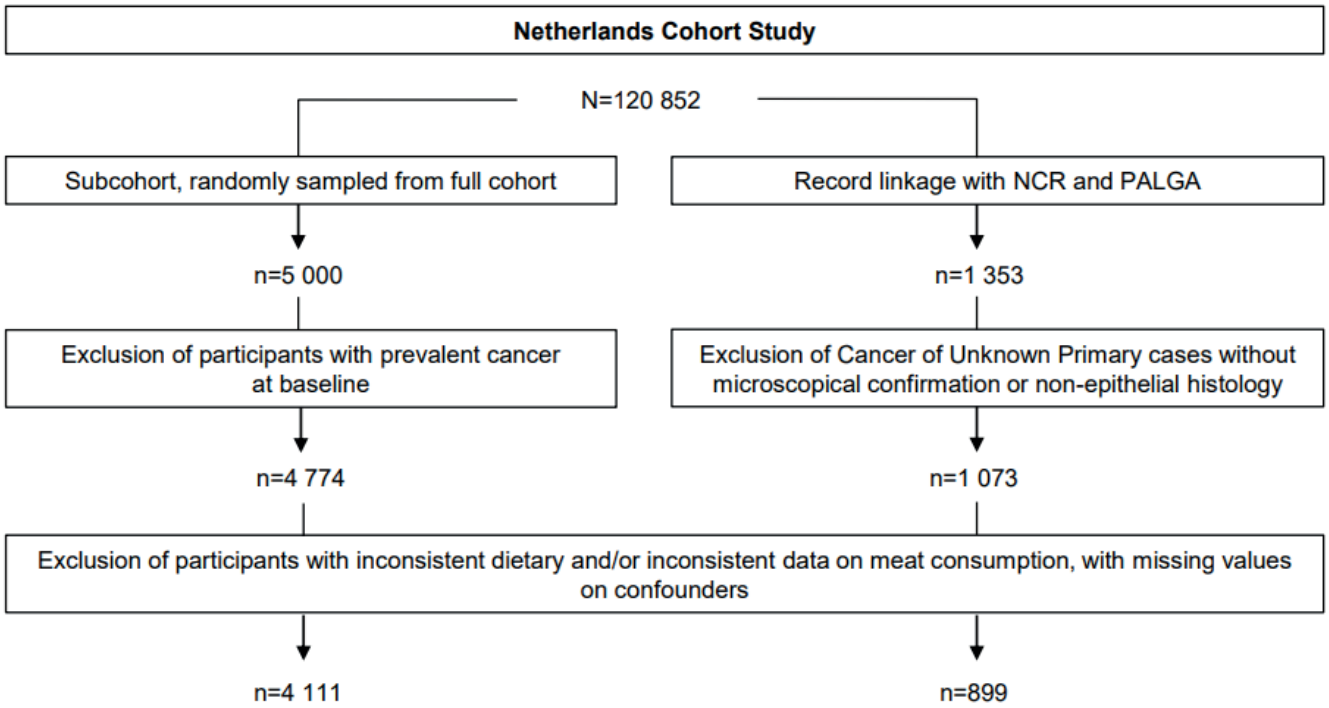


Figure 1 Flow diagram of subcohort members and Cancer of Unknown Primary cases in the Netherlands Cohort Study on whom analyses are based

Results

The statistical analyses of this study are based on 899 incident CUP cases and 4,111 subcohort members with complete and consistent dietary data. CUP cases appeared to consume slightly more red meat (overall), processed meat, and fish than subcohort members (90.8 g/day versus 86.9 g/day & 15.0 g/day versus 13.1 g/day & 14.1 g/day versus 12.9 g/day, respectively) (see Table 1). By contrast, subcohort members ate slightly more poultry than CUP cases (13.5 g/day versus 12.9 g/day). The comparison between CUP cases and subcohort members appeared to be confounded by sex with respect to consuming red meat (overall) and processed meat. Male CUP cases consumed more red meat (overall) than male subcohort members (95.1 g/day versus 93.4 g/day). Female CUP cases consumed more red meat (overall) than female subcohort members (83.5 g/day versus 80.6 g/day). In addition, male CUP cases ate slightly more processed meats than male subcohort members (16.4 g/day versus 15.8 g/day). Female CUP cases also ate more processed meats than female subcohort members (12.7 g/day versus 10.4 g/day). Neither poultry consumption nor fish consumption appeared to be confounded by sex.

Findings of the age- and sex- adjusted analyses were comparable to those of the multivariable adjusted analyses, which were additionally adjusted for alcohol consumption, cigarette smoking variables (status, frequency, duration), and total energy intake. Hence, only the results of the multivariable analyses are described below. In general, we observed no association between red meat (overall) consumption and CUP risk (HR for Q4 vs. Q1: 1.04, 95% CI: 0.83-1.30, P_{trend} = 0.31) (see Table 2). We observed an increased risk between beef consumption and CUP, for which a statistically significant trend was found (HR for Q4 vs. Q1: 1.22, 95% CI: 0.99-1.52, P_{trend} = 0.02). A statistically significant association was also observed between processed meat consumption and CUP risk (HR for Q4 vs. Q1: 1.40, 95% CI: 1.12-1.75, P_{trend} = 0.006). No association was found between poultry consumption and CUP risk (HR for C4 vs. C1: 0.97, 95% CI: 0.79-1.21, P_{trend} = 0.28). For fish consumption, we observed an increased CUP risk, but it was not statistically significant (HR for Q4 vs. Q1: 1.25, 95% CI: 0.99-1.57, P_{trend} = 0.29).

As described above, meat consumption differed markedly between men and women concerning both red meat (overall) and processed meat. Therefore, we stratified the analyses based on sex (see Table 3). For beef consumption and CUP risk in men alone, the association attenuated and the trend was no longer statistically

significant (HR for Q4 vs. Q1: 1.12, 95% CI: 0.85-1.47, P_{trend} = 0.31). Conversely, for beef consumption and CUP risk in women alone, the association became stronger and was statistically significant (HR for Q4 vs. Q1: 1.47, 95% CI: 1.04-2.07, P_{trend} = 0.004). For processed meat consumption and CUP risk in men alone, the association slightly attenuated and was no longer statistically significant (HR for Q4 vs. Q1: 1.33, 95% CI: 0.99-1.79, P_{trend} = 0.15). Yet, the association appeared to be more pronounced in women and remained statistically significant (HR for Q4 vs. Q1: 1.53, 95% CI: 1.08-2.16, P_{trend} = 0.001).

Furthermore, we checked whether there was a potential for residual confounding by cigarette smoking status and the association between meat consumption and CUP risk. We observed that the associations between beef and processed meat consumption and CUP risk increased when comparing current smokers to never smokers in women (data not shown). It should, however, be acknowledged that there were fewer cases available in the categories due to the stratification for both sex and cigarette smoking status. Our observations suggest that residual confounding by cigarette smoking status is unlikely in women.

We observed no multiplicative interactions between sex and the consumption of red meat (overall), beef, pork, minced meat, liver, processed meat, poultry, or fish in relation to CUP risk ($P_{interaction}$ = 0.64, 0.55, 0.22, 0.19, 0.41, 0.52, 0.11, and 0.22, respectively). In addition, no multiplicative interactions were observed between cigarette smoking status and the consumption of red meat (overall), beef, pork, minced meat, liver, processed meat, poultry, or fish in relation to CUP risk ($P_{interaction}$ = 0.27, 0.88, 0.22, 0.56, 0.14, 0.24, 0.88, and 0.80, respectively).

Table 1 Characteristics of Cancer of Unknown Primary cases and subcohort members in the Netherlands Cohort Study

	Subcohort members (n=4111)		Cancer of Unknown Primary cases (n=899)	
Characteristic	**n**	**(%)**	**n**	**(%)**
Age at baseline (years)				
55-59	1605	39.0	276	30.7
60-64	1402	34.1	349	38.8
65-69	1104	26.9	274	30.5
Sex				
Men	2022	49.2	568	63.2
Women	2089	50.8	331	36.8
Red meat consumption (g/day)	86.9 (40.5)		90.8 (40.3)	
Men	93.4 (41.6)		95.1 (41.6)	
Women	80.6 (38.3)		83.5 (37.0)	
Processed meat consumption (g/day)	13.1 (14.6)		15.0 (15.5)	
Men	15.8 (16.7)		16.4 (15.6)	
Women	10.4 (11.6)		12.7 (15.0)	
Poultry consumption (g/day)	13.5 (15.1)		12.9 (14.4)	
Men	13.6 (14.7)		13.3 (15.3)	
Women	13.4 (15.4)		12.2 (12.8)	
Fish consumption (g/day)	12.9 (15.3)		14.1 (17.9)	
Men	14.1 (16.6)		15.5 (20.2)	
Women	11.7 (13.9)		11.7 (12.8)	
Ethanol intake (grams/day) [1]				
Abstainers	975	23.7	170	18.9
<5	1179	28.7	233	25.9
5-<15	938	22.8	204	22.7
15-<30	651	15.8	143	15.9
≥30	368	9.0	149	16.6

Characteristic	Subcohort members (n=4111) n	(%)	Cancer of Unknown Primary cases (n=899) n	(%)
Cigarette smoking status				
Never smokers	1517	36.9	249	27.7
Ex smokers	1479	36.0	317	35.3
Current smokers	1115	27.1	333	37.0
Frequency of cigarette smoking (N/day), mean (SD) [2]	15.7 (10.1)		17.9 (10.4)	
Duration of cigarette smoking (years), mean (SD) [2]	31.8 (12.1)		35.4 (11.7)	
Body Mass Index at baseline (kg/m²), mean (SD)	25.0 (3.1)		24.9 (3.0)	
Non-occupational physical activity (min/day)				
≤30	844	20.8	182	20.5
>30-60	1269	31.2	274	30.9
>60-90	857	21.1	160	18.0
>90	1093	26.9	271	30.6
Level of education (years of education)				
Primary	1154	28.2	237	26.6
Lower vocational	899	22	192	21.6
Secondary and medium vocational	1457	35.6	329	36.9
University and higher vocational	582	14.2	133	14.9
Diabetes				
Yes	140	3.4	36	4.0
First grade family history of cancer				
Yes	1870	45.5	436	48.5

Notes

[1] In consumers only

[2] In users only

Table 2 Hazard ratios and 95% confidence intervals for meat consumption and Cancer of Unknown Primary risk in the Netherlands Cohort Study

	Categorical median (grams per day)		Subcohort members (n=4111)	Cancer of Unknown Primary cases (n=899)					
	Men	Women	Person time at risk (years)	Cases	Age- and sex-adjusted [1]			Multivariable adjusted [2]	
				n	HR	95% CI		HR	95% CI
Red meat (overall)									
Q1	50	41	17 433	205	1	Reference		1	Reference
Q2	79	68	17 338	224	1.14	(0.92-	1.41)	1.11	(0.89- 1.37)
Q3	100	88	17 184	248	1.26	(1.02-	1.55)	1.21	(0.98- 1.49)
Q4	139	125	17 548	222	1.13	(0.91-	1.40)	1.04	(0.83- 1.30)
p for trend [3]					0.08			0.31	
Continuous, 50 grams per day increments			69 503	899	1.08	(0.99-	1.18)	1.05	(0.96- 1.15)
Beef									
Q1	4	3	17 293	199	1	Reference		1	Reference
Q2	16	14	17 065	208	1.05	(0.84-	1.30)	1.03	(0.82- 1.28)
Q3	30	26	17 986	231	1.10	(0.89-	1.36)	1.08	(0.87- 1.34)
Q4	53	47	17 160	261	1.25	(1.01-	1.54)	1.22	(0.99- 1.52)
p for trend [3]					0.01			0.02	
Continuous, 50 grams per day increments			69 503	899	1.21	(1.04-	1.41)	1.21	(1.03- 1.42)
Pork									
Q1	9	5	17 461	214	1	Reference		1	Reference
Q2	28	23	17 308	250	1.21	(0.98-	1.48)	1.14	(0.93- 1.41)
Q3	44	40	17 288	216	1.07	(0.86-	1.32)	1.01	(0.81- 1.25)
Q4	74	65	17 445	219	1.09	(0.88-	1.35)	0.99	(0.79- 1.23)
p for trend [3]					0.65			0.78	
Continuous, 50 grams per day increments			69 503	899	1.03	(0.91-	1.16)	0.98	(0.86- 1.12)
Minced meat									
Q1	3	0	16 932	237	1	Reference		1	Reference
Q2	12	10	17 106	239	1.01	(0.82-	1.24)	1.01	(0.82- 1.24)
Q3	21	18	17 801	207	0.87	(0.70-	1.07)	0.86	(0.70- 1.07)
Q4	38	33	17 664	216	0.90	(0.73-	1.11)	0.91	(0.74- 1.13)
p for trend [3]					0.23			0.22	
Continuous, 50 grams per day increments			69 503	899	0.87	(0.69-	1.09)	0.86	(0.68- 1.09)

	Categorical median (grams per day)		Subcohort members (n=4111)	Cancer of Unknown Primary cases (n=899)					
			Person time at risk (years)	Cases	Age- and sex-adjusted [1]			Multivariable adjusted [2]		
	Men	Women		n	HR	95% CI		HR	95% CI	
Liver										
C1	0	0	44 786	585	1	Reference		1	Reference	
C2	4	3	24 716	314	0.99	(0.84-	1.16)	0.96	(0.82-	1.12)
p for trend [3]					0.92			0.87		
Continuous, 25 grams per day increments			69 503	899	1.02	(0.65-	1.63)	0.96	(0.59-	1.56)
Poultry										
C1	0	0	16 123	203	1	Reference		1	Reference	
C2	5	5	17 045	230	1.12	(0.90-	1.38)	1.13	(0.91-	1.41)
C3	13	13	16 570	227	1.12	(0.90-	1.38)	1.10	(0.88-	1.37)
C4	23	23	19 766	239	0.98	(0.79-	1.21)	0.97	(0.79-	1.21)
p for trend [3]					0.31			0.28		
Continuous, 50 grams per day increments			69 503	899	0.88	(0.68-	1.13)	0.86	(0.66-	1.13)
Processed meat										
Q1	1	0	17 060	201	1	Reference		1	Reference	
Q2	8	4	17 949	216	1.09	(0.88-	1.35)	1.07	(0.86-	1.33)
Q3	16	11	17 095	222	1.16	(0.93-	1.44)	1.14	(0.91-	1.42)
Q4	33	22	17 398	260	1.38	(1.12-	1.70)	1.40	(1.12-	1.75)
p for trend [3]					0.006			0.006		
Continuous, 25 grams per day increments			69 503	899	1.16	(1.04-	1.30)	1.19	(1.05-	1.34)
Fish										
Q1	0	0	19 848	208	1	Reference		1	Reference	
Q2	5	5	15 592	216	1.29	(1.04-	1.59)	1.30	(1.04-	1.61)
Q3	15	15	21 045	286	1.26	(1.03-	1.54)	1.23	(1.00-	1.51)
Q4	32	28	13 018	189	1.32	(1.06-	1.65)	1.25	(0.99-	1.57)
p for trend [3]					0.08			0.29		
Continuous, 25 grams per day increments			69 503	899	1.11	(0.99-	1.24)	1.07	(0.95-	1.20)

Notes

[1] Analyses were adjusted for age at baseline (years) and sex.

[2] Analyses were adjusted for age at baseline (years), sex, alcohol consumption, cigarette smoking status (never/ever), cigarette smoking frequency (continuous; centered), cigarette smoking duration (continuous; centered), and total energy intake (kcal/day).

[3] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

Table 3 Hazard ratios and 95% confidence intervals for meat consumption and Cancer of Unknown Primary risk in the Netherlands Cohort Study, stratified for sex

	Men only					Women only				
	Sub cohort members (n=2022)		Cancer of Unknown Primary cases (n=568)			Sub cohort members (n=2089)		Cancer of Unknown Primary cases (n=331)		
	Categ. median (grams per day)	Person time at risk (years)	Cases	Multivariable adjusted [1]		Categ. median (grams per day)	Person time at risk (years)	Cases	Multivariable adjusted [1]	
			n	HR	95% CI			n	HR	95% CI
Red meat (overall)										
Q1	50	8 143	135	1	Reference	41	9 290	70	1	Reference
Q2	79	7 692	142	1.07	(0.81- 1.42)	68	9 376	82	1.19	(0.85- 1.68)
Q3	100	7 856	146	1.10	(0.83- 1.44)	88	9 328	102	1.43	(1.03- 2.00)
Q4	139	8 222	145	1.01	(0.76- 1.35)	125	9326	77	1.09	(0.77- 1.54)
p for trend [2]				0.70					0.20	
Continuous, 50 grams per day increments		32 182	568	1.02	(0.91- 1.16)		37 320	331	1.10	(0.95- 1.27)
Beef										
Q1	4	7 988	133	1	Reference	3	9 305	66	1	Reference
Q2	16	8 003	134	0.97	(0.73- 1.29)	14	9 062	74	1.15	(0.81- 1.65)
Q3	30	8 116	138	0.97	(0.73- 1.29)	26	9 870	93	1.34	(0.95- 1.88)
Q4	53	8 076	163	1.12	(0.85- 1.47)	47	9 083	98	1.47	(1.04- 2.07)
p for trend [2]				0.31					0.004	
Continuous, 50 grams per day increments		32 182	568	1.11	(0.91- 1.37)		37 320	331	1.41	(1.11- 1.78)
Pork										
Q1	9	8 157	128	1	Reference	5	9 305	86	1	Reference
Q2	28	8 037	151	1.13	(0.85- 1.49)	23	9 271	99	1.16	(0.85- 1.59)
Q3	44	7 787	137	1.07	(0.80- 1.43)	40	9 502	79	0.90	(0.64- 1.25)
Q4	74	8 202	152	1.15	(0.86- 1.53)	65	9 244	67	0.75	(0.53- 1.06)
p for trend [2]				0.49					0.10	
Continuous, 50 grams per day increments		32 182	568	1.06	(0.90- 1.25)		37 320	331	0.83	(0.66- 1.04)
Minced meat										
Q1	3	7 829	162	1	Reference	0	9 103	75	1	Reference
Q2	12	8 049	151	0.92	(0.70- 1.20)	10	9 056	88	1.20	(0.86- 1.68)

Q3	21	8 138	127	0.76	(0.58-	1.00)	18	9 663	80	1.07	(0.76-	1.52)
Q4	38	8 166	128	0.78	(0.59-	1.03)	33	9 498	88	1.19	(0.85-	1.67)
p for trend [2]				0.07						0.53		
Continuous, 50 grams per day increments		32 182	568	0.76	(0.56-	1.03)		37 320	331	1.13	(0.77-	1.64)
Liver												
C1	0	19 695	362	1	Reference		0	25 092	223	1	Reference	
C2	4	12 488	206	0.92	(0.75-	1.12)	3	12 229	108	1.03	(0.80-	1.33)
p for trend [2]				0.88						0.98		
Continuous, 25 grams per day increments		32 182	568	0.95	(0.52-	1.75)		37 320	331	0.99	(0.45-	2.19)
Poultry												
C1	0	7 542	121	1	Reference		0	8 580	82	1	Reference	
C2	5	7 248	151	1.33	(1.00-	1.76)	5	9 797	79	0.87	(0.62-	1.21)
C3	13	8 044	139	1.08	(0.81-	1.44)	13	8 526	88	1.14	(0.82-	1.59)
C4	23	9 348	157	1.07	(0.80-	1.41)	23	10 417	82	0.84	(0.60-	1.17)
p for trend [2]				0.87						0.11		
Continuous, 50 grams per day increments		32 182	568	0.97	(0.67-	1.40)		37 320	331	0.73	(0.50-	1.07)
Processed meat												
Q1	1	8 075	135	1	Reference		0	8 986	66	1	Reference	
Q2	8	7 974	126	0.98	(0.74-	1.30)	4	9 975	90	1.25	(0.88-	1.76)
Q3	16	8 097	149	1.15	(0.87-	1.53)	11	8 998	73	1.12	(0.78-	1.61)
Q4	33	8 036	158	1.33	(0.99-	1.79)	22	9 362	102	1.53	(1.08-	2.16)
p for trend [2]				0.15						0.001		
Continuous, 25 grams per day increments		32 182	568	1.11	(0.96-	1.28)		37 320	331	1.50	(1.17-	1.93)
Fish												
Q1	0	8 231	119	1	Reference		0	11 617	89	1	Reference	
Q2	5	7 256	144	1.36	(1.02-	1.81)	5	8 336	72	1.18	(0.84-	1.66)
Q3	15	10 054	169	1.12	(0.85-	1.48)	15	10 991	117	1.43	(1.05-	1.95)
Q4	32	6 642	136	1.31	(0.98-	1.76)	28	6 377	53	1.10	(0.76-	1.61)
p for trend [2]				0.22						0.99		
Continuous, 25 grams per day increments		32 182	568	1.10	(0.95-	1.27)		37 320	331	1.00	(0.82-	1.21)

Notes

[1] Analyses were adjusted for age at baseline (years), alcohol consumption, cigarette smoking status (never/ever), cigarette smoking frequency (continuous; centered), cigarette smoking duration (continuous; centered), and total energy intake (kcal/day).

[2] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

Results from the first sensitivity analysis with restriction to histologically verified CUP cases alone, for whom extended diagnostic methods are more expected (compared to cytologically verified CUP cases), indicate that the findings are similar to those of the overall multivariable analyses except for beef consumption and CUP risk (HR for Q4 vs. Q1: 1.16, 95% CI: 0.91-1.49, P_{trend} = 0.21), possibly due to fewer cases. We presume that the results of the overall multivariable analyses represent CUP cases with or without an extensive diagnostic work-up. Our secondary sensitivity analysis, in which the first two years of follow-up were excluded so as to check for potential reverse causality bias, also demonstrate similar findings to those observed in the complete analysis (data not shown). In our third sensitivity analysis, after splitting the follow-up time to compare the first ten years of follow-up to the last ten years of follow-up, we observed that the association between beef consumption and CUP risk was the highest in the first ten years of follow-up, whereas it attenuated in the last ten years of follow-up. On the other hand, for processed meat consumption and CUP risk, no association was found in the first ten years of follow-up, while there was a positive statistically significant association in the last ten years of follow-up.

Discussion

In this detailed investigation of meat consumption and CUP risk, we found that beef and processed meat consumption were positively associated with the development of CUP in women. We found a non-significant positive association between processed meat consumption and CUP risk in men. In contrast, no associations were observed between red meat (overall), poultry, or fish consumption and CUP risk. We observed no multiplicative interactions between sex or cigarette smoking status and meat consumption and CUP risk.

To the best of our knowledge, only one study has previously investigated the relationship between red meat and processed meat and CUP risk. The abovementioned Australian cohort study compared 327 incident CUP cases to two sets of controls (3:1) that were randomly selected using incidence density sampling with replacement. Their study found no relation between red meat consumption and CUP risk; it used the usual number of servings as >3 red meats/week compared to <3 red meats/week for dichotomous comparisons [11]. For processed meat consumption and CUP, its authors observed an increased risk when comparing the usual number of servings as >3 processed meats/week compared to <3 processed meats/week, although the association was not statistically significant [11]. In the

NLCS, by contrast, we have investigated the association between meat consumption and CUP risk in greater detail by assessing combined groups of meats such as red meat, processed meat, poultry, and fish, as well as individual meat items. We have found that beef and processed meat consumption are significantly associated with an increased CUP risk, but that red meat (overall), poultry, and fish consumption do not appear to be associated with CUP risk. Consequently, while our study confirms the findings of the Australian cohort study in indicating no association between red meat (overall) consumption and CUP risk, we do observe an association between beef and processed meat consumption and CUP risk [11]. The consumption of red and processed meat has been linked to colorectal cancer in previous epidemiological studies (probable increasing risk and convincing increasing risk, respectively) [26]. It also known that colorectal cancer predominantly metastasises to the liver via the portal circulation [27], therefore, we have conducted an additional analysis to study whether the association between meat consumption is stronger in CUP patients with metastases located in the liver. We found the association between processed meat consumption and CUP risk in patients with a liver metastasis to be increased (per 25 grams per day increment HR: 1.34, 95% CI: 1.14-1.58, P_{trend} = 0.001) compared to the result of the overall analysis (per 25 grams per day increment HR: 1.19, 95% CI: 1.05-1.34, P_{trend} = 0.006). In addition, based on data obtained from the NCR, 36.1% of the primary tumours that metastasised to the liver, originated in the colorectum (ICD-O-3 C18-C20) between 1986 and 2006 in the Dutch population. In line with the results of our analysis, it is thus plausible that in a considerable number of CUP patients with a liver metastasis, the primary tumour origin is the colorectum. Furthermore, we have checked the potential of residual confounding by cigarette smoking status. Despite studying fewer cases in the categories of interest due to stratification based on sex and cigarette smoking status, the association between beef and processed meat consumption did not differ greatly between the strata (never, ex, current smokers) in women, thereby hinting that the potential of residual confounding is unlikely. We have also checked whether splitting the follow-up time had an influence on the association between meat consumption and CUP risk. We observed that the association between beef consumption and CUP risk was highest in the first ten years of follow-up, whereas it attenuated in the last ten years of follow-up. For processed meat consumption and CUP risk, no association was found in the first ten years of follow-up, while there was a positive statistically significant association in the last ten years of follow-up. An indication for these findings might be that there is a shorter latency period for beef consumption and a relatively longer

latency period for processed meat consumption, or that it concerns a chance finding as there were fewer cases available due to splitting the follow-up time. Therefore, more studies would be needed to investigate such conclusions.

As briefly presented in the introduction, scientific evidence has already revealed associations between red meat intake and processed meat intake and the development of specific cancers, though the associations are less consistent concerning poultry and fish consumption and carcinogenesis [13,17]. As we have demonstrated here, however, there does appear to be a discernible connection between the consumption of beef and processed meats and the development of CUP.

Strengths and limitations

Important strengths of this study are its prospective cohort design, large sample size of 120,852 participants, large number of incident CUP cases, and the detailed availability of exposure and confounder data. Moreover, completeness of record linkage with the NCR and PALGA for cancer follow-up was at least 96%, which minimizes selection bias [28]. Vital status follow-up was complete for almost 100% after 20.3 years. Details on incident CUP cases were obtained from the NCR and included specific information from both pathology reports and clinical reports [29]. In addition, we could access the pathology excerpts and thus check whether the cytological and/or histological confirmed cases had been correctly categorised in the data provided by the NCR. The NCR registry clerks applied uniform coding rules when entering data based on medical files.

There are certain limitations that should be acknowledged. Exposure data on meat consumption were only measured at baseline in 1986, so participants may have changed their dietary habits after having completed the questionnaire, which could result in bias due to misclassification. The questionnaire was tested, however, both for validity and reproducibility purposes and appeared to be representative for dietary habits over a period of at least five years [22,23]. In addition, this potential bias should be non-differential between CUP cases and subcohort members.

Conclusions

Beef and processed meat consumption appear to be positively associated with CUP risk in women. Similarly, a positive association was found between processed meat consumption and CUP risk in men, although it was not statistically significant. We found no associations between red meat (overall), poultry, or fish consumption and CUP risk.

References

1. Loffler H, Kramer A (2016) Introduction. In: Loffler H, Kramer A (eds) Cancer of Unknown Primary Springer International Publishing Switzerland. doi:10.1007/978-3-319-22581-4
2. Schroten-Loef C, Verhoeven R, de Hingh I, van de Wouw A, van Laarhoven H, Lemmens V (2018) Unknown primary carcinoma in the Netherlands: decrease in incidence and survival times remain poor between 2000 and 2012. Eur J Cancer 101:77-86
3. Comprehensive Cancer Centre the Netherlands (2020) Incidentie Primaire Tumor Onbekend. Integraal kankercentrum Nederland. https://iknl.nl/kankersoorten/primaire-tumor-onbekend/registratie/incidentie].
4. Meijer L, De Peuter R, Van der Zwan JM, Loef C (2020) Primaire Tumor Onbekend. Integraal kankercentrum Nederland. https://www.iknl.nl/getmedia/ea952e4d-9af7-4d59-9f03-865c858bc0cd/Rapport_PrimaireTumorOnbekend_IKNL_2020.pdf].
5. Pavlidis N, Pentheroudakis G (2012) Cancer of unknown primary site. Lancet 379 (9824):1428-1435. doi:10.1016/S0140-6736(11)61178-1
6. Hemminki K (2016) Epidemiology, Risk Factors, and Survival in CUP: Pointers to Disease Mechanisms. In: Cancer of Unknown Primary. Springer, pp 5-25
7. Rassy E, Pavlidis N (2020) Progress in refining the clinical management of cancer of unknown primary in the molecular era. J Nature Reviews Clinical Oncology 17 (9):541-554
8. Fizazi K, Greco FA, Pavlidis N, Daugaard G, Oien K, Pentheroudakis G, Committee EG (2015) Cancers of unknown primary site: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 26 Suppl 5 (Suppl 5):v133-138. doi:10.1093/annonc/mdv305
9. Hemminki K, Chen B, Melander O, Manjer J, Hallmans G, Hemminki A (2015) Smoking and body mass index as risk factors for subtypes of cancer of unknown primary. Int J Cancer 136 (1):246-247. doi:10.1002/ijc.28969
10. Kaaks R, Sookthai D, Hemminki K, Kramer A, Boeing H, Wirfalt E, Weiderpass E, Overvad K, Tjonneland A, Olsen A, Peeters PH, Bueno-de-Mesquita HB, Panico S, Pala V, Vineis P, Quiros JR, Ardanaz E, Sanchez MJ, Chirlaque MD, Larranaga N, Brennan P, Trichopoulos D, Trichopoulou A, Lagiou P, Hallmans G, Khaw KT, Key TJ, Riboli E, Canzian F (2014) Risk factors for cancers of unknown primary site: Results from the prospective EPIC cohort. Int J Cancer 135 (10):2475-2481. doi:10.1002/ijc.28874
11. Vajdic C, Perez-Concha O, Dobbins T, Ward RL, Schaffer AL, van Leeuwen M, Rhee JJ, Laaksonen M, Craigen G, Pearson S-A (2019) Demographic, social and lifestyle risk factors for cancer registry-notified cancer of unknown primary site (CUP). Cancer Epidemiol 60:156-161
12. Hermans KE, van den Brandt PA, Loef C, Jansen RL, Schouten LJ (2021) Alcohol consumption, cigarette smoking and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study. Int J Cancer
13. Bouvard V, Loomis D, Guyton KZ, Grosse Y, Ghissassi FE, Benbrahim-Tallaa L, Guha N, Mattock H, Straif K, Corpet DJTLO (2015) Carcinogenicity of consumption of red and processed meat. 16 (16):1599-1600
14. Ferguson L (2010) Meat and cancer. Meat Science 84 (2):308-313
15. Alomirah H, Al-Zenki S, Al-Hooti S, Zaghloul S, Sawaya W, Ahmed N, Kannan KJFc (2011) Concentrations and dietary exposure to polycyclic aromatic hydrocarbons (PAHs) from grilled and smoked foods. 22 (12):2028-2035

16. Daniel CR, Cross AJ, Graubard BI, Hollenbeck AR, Park Y, Sinha R (2011) Prospective investigation of poultry and fish intake in relation to cancer risk. Cancer prevention research 4 (11):1903-1911
17. World Cancer Research Fund/American Institute for Cancer Research (2018) Meat, fish and dairy products and the risk of cancer. https://www.wcrf.org/dietandcancer.
18. Van den Brandt PA, Goldbohm RA, van 't Veer P, Volovics A, Hermus RJ, Sturmans F (1990) A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol 43 (3):285-295
19. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM (1990) Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol 19 (3):553-558
20. Goldbohm RA, van den Brandt PA, van 't Veer P, Brants HA, Dorant E, Sturmans F, Hermus RJ (1994) A prospective cohort study on the relation between meat consumption and the risk of colon cancer. Cancer Res 54 (3):718-723
21. Van den Brandt P (2019) Red meat, processed meat, and other dietary protein sources and risk of overall and cause-specific mortality in The Netherlands Cohort Study. Eur J Epidemiol 34 (4):351-369
22. Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, Hermus RJ (1994) Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin Nutr 48 (4):253-265
23. Goldbohm RA, van 't Veer P, van den Brandt PA, van 't Hof MA, Brants HA, Sturmans F, Hermus RJ (1995) Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. Eur J Clin Nutr 49 (6):420-429
24. Barlow WE (1994) Robust variance estimation for the case-cohort design. Biometrics:1064-1072
25. Lin DY, Wei L-J (1989) The robust inference for the Cox proportional hazards model. J Am Stat Assoc 84 (408):1074-1078
26. World Cancer Research Fund/American Institute for Cancer Research (2018) Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. Continuous Update Project Expert Report 2018.
27. Riihimäki M, Hemminki A, Sundquist J, Hemminki K (2016) Patterns of metastasis in colon and rectal cancer. J Scientific reports 6 (1):1-9
28. Goldbohm RA, Van den Brandt PA, Dorant E (1994) Estimation of the coverage of Dutch municipalities by cancer registries and PALGA based on hospital discharge data. Tijdschr Soc Gezondheidsz 72:80-84
29. Van der Sanden G, Coebergh J-W, Schouten L, Visser O, Van Leeuwen F (1995) Cancer incidence in the Netherlands in 1989 and 1990: First results of the nationwide Netherlands Cancer Registry. Eur J Cancer 31 (11):1822-1829

CHAPTER 6

FAMILY HISTORY OF CANCER IN FIRST DEGREE RELATIVES AND CANCER OF UNKNOWN PRIMARY RISK

Alexander L.R. Grewcock, Karlijn E.P.E. Hermans, Matty P. Weijenberg, Piet A. van den Brandt, Caroline Loef, Rob L. H. Jansen, and Leo J. Schouten

Eur. J. Cancer Care. 2021;1-9

Abstract

Objective: Cancer of Unknown Primary (CUP) refers to the presence of metastatic lesions, with no identifiable primary site during the patient's lifetime. Poor survival and lack of available treatment highlight the need to identify potential CUP risk factors. We investigated whether a family history of cancer is associated with increased CUP risk.

Methods: We performed a case cohort analysis using data from the Netherlands Cohort Study, which included a total of 963 CUP cases and 4,288 subcohort members. A Cox Proportional Hazards Regression was used to compare CUP risk in participants who reported to have a family member with cancer to those who did not, whilst adjusting for confounders.

Results: In general, we observed no increased CUP risk in those who reported a family history of cancer. CUP risk appeared slightly increased in those who reported cancer in a sibling (HR: 1.16, 95% CI: 0.97-1.38), especially in those with a sister with cancer compared to those without (HR: 1.23, 95% CI: 0.99-1.53), although these findings are not statistically significant.

Conclusion: Having a family history of cancer is not an independent risk factor of CUP.

Introduction

Cancer of Unknown Primary (CUP) refers to the presence of metastatic lesions in a patient without an identifiable primary site (National Institute for Health and Care Excellence, 2010). Globally, CUP incidence has been decreasing. This decrease may be partly explained by improved imaging techniques and molecular investigation(s) used to identify primary tumour sites (Rassy & Pavlidis, 2019). It is difficult to determine the true international incidence and prevalence of CUP, centres define CUP differently and definitions have varied over time within centres. Nevertheless, in the Netherlands, CUP accounted for approximately 1,300 patients in 2018 (Comprehensive Cancer Centre the Netherlands, 2020).

Despite advances in diagnostics leading to identification of primary sites in patients that would previously have been classified as CUP patients, the limited improvement in treatment means CUP remains difficult to treat. Therefore, the prognosis for most CUP patients is notoriously poor, with a median survival of around 2 months (Schroten-Loef et al., 2018). The limited opportunity for curative and life-prolonging treatment highlights the need for a preventative approach to managing CUP (Rassy, Assi, & Pavlidis, 2020). Such approaches require identification of risk factors as well as identification of people most at risk, which is challenging given that CUP aetiology studies are relatively understudied.

Demographic factors appear to be important for CUP risk, since increased CUP risks are seen both in women and with increasing age (Luke et al., 2008). Studies in younger patients demonstrate higher rates of CUP incidence in metropolitan areas with lower socio-economic status. A higher prevalence of potential risk factors and reduced access to healthcare, and/or overdiagnosis of CUP as a result of poorer access to diagnostic facilities that specifically identify primary tumours could explain these findings (Pavlidis, Rassy, & Smith-Gagen, 2020). Additionally, modifiable lifestyle related risk factors have been highlighted as influential. For instance, CUP is associated with cigarette smoking (Hemminki, Chen, et al., 2015; Hermans, van den Brandt, Loef, Jansen, & Schouten, 2021; Kaaks et al., 2014; Vajdic et al., 2019). Similarly, alcohol consumption is also associated with CUP risk in a dose-response relationship (Hermans et al., 2021). A weaker association was found for waist circumference which was no longer statistically significant after adjusting for confounders (Kaaks et al., 2014).

Some evidence shows that CUP is associated with a multitude of pre-existing health conditions. In an Australian population, CUP patients were found to be

more likely to suffer with diabetes and a pre-existing cancer diagnosis (Vajdic et al., 2019). This was also seen in a Swedish population where CUP was associated with diabetes and various autoimmune disorders (Hemminki, Försti, Sundquist, & Li, 2016; Hemminki, Sundquist, Sundquist, & Ji, 2015).

The lack of studies that investigate the associations between CUP and modifiable and demographic characteristics makes it difficult to draw firm conclusions on which factors increase CUP risk. This is also the case for the possible familial aspects of CUP. The possible role of genetic susceptibility and shared environmental factors contributing to increase CUP risk, is hinted at by the extensive evidence for clustering of cancer within families across anatomical sites (Hemminki, Bevier, Sundquist, & Hemminki, 2012; Hemminki, Ji, Sundquist, & Shu, 2011; Zeegers, Schouten, Goldbohm, & van den Brandt, 2008).

This propensity for familial clustering also appears to be a trait of CUP, as familial clustering was demonstrated in a study using the Swedish Family Cancer Database, which found CUP patients were more likely to have a sibling with CUP. Moreover, patients who had a diagnosis of lung, liver, kidney, pancreatic, ovarian, or colorectal cancer were also more likely to have a family member diagnosed with CUP. The same authors redemonstrated these associations using an updated version of the database (Hemminki et al., 2011; Hemminki, Sundquist, Sundquist, Hemminki, & Ji, 2016). This finding is supported by evidence from a nested case control study in a Utah population which similarly found an increased CUP risk, as well as increased risk of lung and pancreatic cancer, myeloma, and non-Hodgkin lymphoma in family members of CUP patients compared to relatives of population controls without CUP (Samadder et al., 2016). Hemminki et al. (2012) examined the association between the anatomical site of cancer in a family member and the risk of metastasis of CUP at that same site. The strongest significant associations were seen for lung, pancreatic, and ovarian cancer, suggesting that the location of the hidden primary in CUP patients may coincide with the anatomical site of cancer in their family members (Hemminki et al., 2012).

These findings imply that CUP may have a familial component, yet the number of studies is small and the studies are limited in terms of variety of populations and the study designs applied. Therefore, in the present study we examined the association between cancer in family members (both overall and in specific relatives) and CUP risk as well as the association between cancer in family members at specific anatomical sites and CUP risk. In order to do so, we formulated the following research

questions: 1) What is the association between a family history of any cancer in first degree relatives and CUP risk? and 2) What is the association between a family history of cancer in first degree relatives at specific anatomical sites and CUP risk?

Methods

Design and study population

The NLCS is a prospective cohort study which started in 1986. Its primary aim was to investigate associations between diet and cancer. The design and methods used in the NLCS are described in detail elsewhere (Van den Brandt, Goldbohm, et al., 1990). A total of 120,852 participants aged 55-69 were sampled from 204 Dutch municipalities. Key demographic variables were extracted from municipal population registries. Participants were asked to complete a baseline questionnaire which entailed detailed information regarding diet and other cancer related risk factors. The case-cohort design was applied for increased efficiency of data processing and analyses. Therefore, a subcohort of 5,000 participants was used to estimate both the person-years at risk accumulated, and the characteristics of the full cohort. The subcohort comprises a randomly selected group of participants at baseline, in whom CUP cases can occur (Barlow, Ichikawa, Rosner, & Izumi, 1999). Participants with a prevalent diagnosis of cancer at recruitment were excluded, unless that diagnosis was skin cancer.

Outcome measure

For this study, CUP cases are patients with either a histologically and/ or cytologically confirmed epithelial malignancy with no identifiable primary site during the patient's lifetime (ICD-O-3: M-8000 - M8570). With the focus on epithelial malignancies, CUP cases who had a histology of sarcomas, lymphomas, mesotheliomas, and melanomas were not considered.

Follow-up

CUP cases were identified from the total cohort of the NLCS during a follow up period of 20.3 years using record linkage to the Netherlands Cancer Registry (NCR) and the Dutch Pathology Registry (PALGA) (Van den Brandt, Schouten, Goldbohm, Dorant, & Hunen, 1990). A total of 963 CUP cases and 4,288 subcohort members were available for analyses after excluding participants with missing data for variables used in the multivariable model.

Questionnaire Data

Data was obtained through a self-administered questionnaire that included detailed questions on dietary information, and other cancer risk factors such as smoking, alcohol consumption, history of cancer and comorbidities. With respect to family history of cancer, participants were asked whether they had a brother, sister, or parent who had cancer. Participants who responded yes were then asked to document the relative affected, the type of cancer, the age at diagnosis, as well as the relative's current age or age of death if applicable. Participants were asked to give information about the number of siblings they had, and if applicable, their year and cause of death. The questionnaire also included questions on smoking behaviour, which was measured based on smoking status (never, former or current smokers), smoking duration (number of years) and smoking frequency (cigarettes per day). The questionnaire also addressed alcohol consumption, most notably the number of alcoholic drinks that had been consumed in the previous week (in glasses), which represented average alcohol consumption in ten grams per day increments. BMI (kg/m^2) was calculated using self-reported height (cm) and weight (kg) at baseline. Participants were asked to state their highest level of education achieved, to represent socioeconomic status. Diabetes status was asked to indicate whether the participant had self-reported a doctor's diagnosis of diabetes in the questionnaire. For non-occupational physical activity (gardening, cycling and walking, and sports/physical exercise), participants could report their activity value, which was summed into a total non-occupational physical activity value.

Statistical Analysis

Characteristics of CUP cases and subcohort members were compared based on the variables of interest. Frequencies and percentages were used for categorical variables, with means and standard deviations for continuous variables. Cox Proportional Hazards Regression was used for case-cohort analyses. Cases derived from the full cohort and the person-time-at-risk for the cohort was calculated using the subcohort. CUP risk was modelled against a family history of cancer to produce hazard ratios (HRs) and 95% confidence intervals (CIs). CUP risk was assessed in participants with any first degree relative with cancer, specifically in siblings or parents as well as discordant anatomical sites. To perform such analyses, three variables were created. The first binary variable compared participants with at least one family member (either a sibling or parent) with cancer to participants with no reported family members with cancer. A binary variable was created to represent specific first-degree relatives including

brothers, sisters, fathers and mothers individually. A separate variable was created both for brothers and sisters to account for the difference in biological sex, a factor which has been demonstrated to influence CUP risk. Similarly, a binary variable was used to compare participants with at least one parent affected with cancer against participants with no parents affected. The CUP risk in participants who reported a family history of cancer at specific sites was also analysed. This analysis was done for breast, ovarian, endometrial, bowel, stomach, lung, kidney, prostate, bladder, pancreas, head and neck, leukaemia, and lymphoma, as it has been shown that family members of patients with such cancers are at an increased CUP risk. Here, binary variables were used to indicate presence or absence of this cancer in the family history.

Age, sex, smoking, and alcohol consumption were considered as predefined confounders and were used in all statistical models, as these factors have been demonstrated to be associated with CUP (Hemminki, Chen, et al., 2015; Hermans et al., 2021; Kaaks et al., 2014; Vajdic et al., 2019). Potential confounders (BMI, socio-economic status, physical activity, and diabetes status) were evaluated using the backward elimination procedure. A variable was considered a confounder if it introduced a greater than 10% change in the HRs once it was removed. Accordingly, none of the potential confounders were included in the final model. Once the variables and interaction terms had been established, CUP was modelled against family history of cancer overall, in siblings and in parents separately. Lastly, CUP was modelled against family history of cancer in discordant anatomical sites of the family members. Scaled Schoenfeld residuals were used to test for the proportional hazards assumption (Lin & Wei, 1989). Log minus log plots were visually inspected for confirmation. If the assumption was deemed to be violated, this was managed by including a time varying covariate (TVC) for the variable at which the violation occurred. Consequently, we added a TVC for age in the age-sex adjusted analysis and for cigarette smoking status and cigarette smoking duration in multivariable analyses. Standard errors were calculated using the robust Huber-White sandwich estimator to account for additional variance introduced by sampling from the full cohort. The Wald test was used to test for multiplicative interaction between age and family history of cancer, sex and family history of cancer, and smoking and family history of cancer. All analyses were conducted using the sixteenth edition of Stata. *P* values below 0.05 indicate statistical significance.

A sensitivity analysis was performed by restricting the analysis to the histologically verified cases of CUP only, as these participants were more likely to have undergone extensive diagnostic investigations before a diagnosis was made. Also, these participants were more likely to meet to more stringent CUP case definitions,

such as those given by NICE (National Institute for Health and Care Excellence, 2010). CUP cases that had been confirmed cytologically but not histologically were excluded from this part of the analysis.

Results

A total of 963 CUP cases and 4,288 subcohort members were available in our multivariable models. The majority of CUP cases were male (62.6%) which differs substantially from the distribution seen in the subcohort (49.2%) (see Table 1). On average, cases were a year older than subcohort members (62 years old and 61 years old, respectively). A greater proportion of cases were current cigarette smokers (37.8%) compared to the subcohort (27.6%). A greater frequency and duration of cigarette smoking was seen amongst smokers in cases compared to smokers in the subcohort. Average alcohol consumption (in grams) was also higher in cases compared to the subcohort, with 14 and 10 grams consumed per day, respectively. A slightly higher proportion of cases reported a family history of cancer in at least one first degree relative (47.7%) compared to the subcohort (45.4%).

Participants who had at least one family member with a history of cancer were not at an increased CUP risk (multivariable adjusted HR: 1.10, 95% CI: 0.95-1.27) compared to participants without (see Table 2). An age-stratified analysis was conducted to obtain age category specific hazard ratios. CUP risk was slightly increased in those aged 60-64 years old (multivariable adjusted HR: 1.27, 95% CI: 1.01-1.61) with a family history of cancer in any relative compared to participants of the same age with no family history of cancer. In terms of siblings and parents, a slightly increased CUP risk was observed in participants with at least one sibling with a history of cancer (multivariable adjusted HR: 1.16, 95% CI: 0.97-1.38) compared to those without, though this was not statistically significant. Multivariable adjusted estimates for parents did not reveal a significant association (HR: 1.02, 95% CI: 0.88-1.19). When mutually adjusting for both siblings and parents, these estimates did not change notably, compared to sibling and parent only analyses. We further adjusted for the number of brothers and sisters the participants had, but this did not alter estimates either. With respect to specific first-degree relatives, a slightly increased CUP risk was observed in participants with a family history of cancer in a sister (multivariable adjusted HR: 1.23, 95% CI: 0.99-1.53), though this was not statistically significant. No association was found in those with a brother with a family history of cancer. Similarly, CUP risk was not increased in those with a family history of cancer in a father compared to those without, nor was the risk increased in those with a family history of cancer in a mother.

Table 1 General characteristics of Cancer of Unknown Primary cases and subcohort members in the Netherlands Cohort Study

	Subcohort members (n=4288)		Cancer of Unknown Primary cases (n=963)	
Exposure variables and potential confounders	**n**	**(%)**	**n**	**(%)**
Age at baseline (years)				
55-59	1164	38.8	288	29.9
60-64	1461	34.1	372	38.6
65-69	1163	27.1	303	31.5
Sex				
Male	2110	49.2	603	62.6
Female	2178	50.8	360	37.4
Family history of cancer				
Yes	1945	45.4	459	47.7
Cigarette smoking status				
Never smokers	1584	36.9	265	27.5
Ex smokers	1521	35.5	334	34.7
Current smokers	1183	27.6	364	37.8
Frequency of cigarette smoking (N/day), mean (SD) [1]	15.7 (10.1)		17.8 (10.4)	
Duration of cigarette smoking (years), mean (SD) [1]	31.9 (12.1)		35.6 (11.6)	
Ethanol intake (grams/day) [2]				
Abstainers	1024	23.9	186	19.3
<5	1228	28.6	247	25.7
5-<15	979	22.8	217	22.5
15-<30	672	15.7	153	15.9
≥30	385	9.0	160	16.6
BMI (kg/m²) at baseline, mean (SD)	25.0 (3.1)		25.0 (3.0)	
Non-occupational physical activity (min/day)				
≤30	908	21.5	204	21.5
>30-60	1318	31.2	291	30.6
>60-90	879	20.8	170	17.9
>90	1122	26.5	285	30.0

	Subcohort members (n=4288)		Cancer of Unknown Primary cases (n=963)	
Exposure variables and potential confounders	**n**	**(%)**	**n**	**(%)**
Level of education (years of education)				
Primary	1257	29.5	271	28.5
Lower vocational	937	22.0	204	21.4
Secondary and medium vocational	1483	34.8	341	35.8
University and higher vocational	590	13.8	136	14.3
Diabetes				
Yes	153	3.6	39	4.1

Notes

[1] In users only

[2] In consumers only

Table 2 Hazard ratios and 95% confidence intervals for Cancer of Unknown Primary risk in participants with family history of cancer in specific relatives in the Netherlands Cohort Study

	Subcohort members	Cancer of Unknown Primary cases				
	Person time at risk (years)	Cases	Age and sex - adjusted [1]		Multivariable adjusted [2]	
First degree relative		**n**	**HR**	**95% CI**	**HR**	**95% CI**
Age at baseline (all ages)						
No	39 347	504	1	Reference	1	Reference
Yes	32 995	459	1.09	(0.94- 1.26)	1.10	(0.95- 1.27)
Age at baseline (ages 55-59)						
No	16 773	149	1	Reference	1	Reference
Yes	13 552	139	1.17	(0.91- 1.50)	1.18	(0.91- 1.52)
Age at baseline (ages 60-64)						
No	13 468	184	1	Reference	1	Reference
Yes	10 919	188	1.27	(1.00- 1.60)	1.27	(1.01- 1.61)

First degree relative		Subcohort members	Cancer of Unknown Primary cases				
		Person time at risk (years)	Cases	Age and sex - adjusted [1]		Multivariable adjusted [2]	
			n	HR	95% CI	HR	95% CI
Age at baseline (ages 65-69)							
	No	9 106	171	1	Reference	1	Reference
	Yes	8 524	132	0.85	(0.66- 1.11)	0.87	(0.67- 1.13)
Siblings							
	No	58 179	744	1	Reference	1	Reference
	Yes	14 163	219	1.17	(0.99- 1.40)	1.16	(0.97- 1.38)
Parents							
	No	48 397	648	1	Reference	1	Reference
	Yes	23 945	315	1.00	(0.86- 1.17)	1.02	(0.88- 1.19)
Sisters							
	No	64 138	828	1	Reference	1	Reference
	Yes	8 204	135	1.24	(1.00- 1.54)	1.23	(0.99- 1.53)
Brothers							
	No	65 049	852	1	Reference	1	Reference
	Yes	7 293	111	1.10	(0.88- 1.38)	1.07	(0.85- 1.35)
Mothers							
	No	60 069	797	1	Reference	1	Reference
	Yes	12 273	166	1.01	(0.83- 1.22)	1.04	(0.86- 1.26)
Fathers							
	No	58 135	774	1	Reference	1	Reference
	Yes	14 207	189	1.03	(0.86- 1.23)	1.04	(0.87- 1.25)

Notes

[1] Analyses were adjusted for age at baseline (years) and sex, and age as a time-varying covariate.

[2] Multivariable analyses were adjusted for age at baseline (years), sex, alcohol consumption (g ethanol/day), current cigarette smoking, cigarette smoking frequency (N/day; continuous; centered), cigarette smoking duration (years; continuous; centered), and cigarette smoking status (never/ever), cigarette smoking duration (continuous; centered) as time-varying covariates.

CUP was not associated with family history of cancer of breast, ovarian, endometrial, bowel, stomach, lung, prostate, bladder, pancreas, head and neck, lymphoma, and/or leukaemia (see Table 3). However, CUP risk appeared to be reduced in those who reported a family history of kidney cancer (multivariable adjusted HR: 0.27, 95% CI: 0.08-0.90), though only three CUP cases reported a family history of kidney cancer.

A total of 687 CUP cases and 4,288 subcohort members were available when the analysis was restricted to histologically verified cases alone. The results of this analysis did not differ markedly from the unrestricted analyses with the exception of the association seen for kidney cancer (data not shown). For kidney cancer, CUP risk remained to be reduced, but it was no longer statistically significant. No multiplicative interaction was detected between age and family history of cancer, between sex and family history of cancer, nor between smoking status and family history of cancer.

Discussion

In this prospective cohort study, having a family history of cancer is not an independent risk factor of CUP. The only consistent association observed was a moderately increased CUP risk in participants who reported a sibling with cancer compared to those who did not. An increased CUP risk was also found in sisters with cancer. However, the association seen for both siblings and sisters were not statistically significant.

Previous studies have investigated CUP risk in relatives of the proband whilst this study has investigated risk in the proband. A cohort study using the Swedish Family-Cancer Database examined CUP risk in family members of patients with various cancers. It demonstrated that people with kidney, lung, and colorectal cancers had higher CUP risks in relatives (Hemminki et al., 2011). This association was stronger for siblings than for parents. This evidence was supported by similar results when the study was repeated using an updated version of the database by the same authors (Hemminki et al., 2011; Hemminki, Sundquist, et al., 2016). Similarly, a nested case control study of an American population (Utah) found an elevated CUP risk in family members of lung, pancreatic, myeloma and non-Hodgkin lymphoma patients compared to relatives of population controls without CUP (Samadder et al., 2016). These three studies were, however, unable to adjust for confounders. To provide further evidence and examination of the family history-CUP association,

we investigated whether this association is present in the opposite direction to previous investigations (Hemminki et al., 2011; Hemminki, Sundquist, et al., 2016; Kaaks et al., 2014; Samadder et al., 2016), by assessing whether CUP risk is increased by the presence of cancer in family members. Extrapolating from the associations seen in these previous studies, we expected CUP risk to be elevated in those with a family history of cancer compared to those without, particularly at the specific cancer sites mentioned above. We observed slightly increased CUP risk in those who reported a sibling with any cancer, but not in parents. This association appears to be accounted for by the increased CUP risk that we observed in participants who reported to have sisters with a diagnosis of cancer compared to participants who did not. In general, the association appears to be comparable with evidence from the Swedish cohort study, in which an increased CUP risk was observed in siblings of patients with cancer at many different anatomical sites. Associations between siblings partly point towards lifestyle-related factors, such as smoking behaviour and alcohol consumption, which may be more similar between siblings, rather than between parents and children.

The findings of the NLCS are inconsistent with the considerable associations observed between CUP risk and discordant cancer sites in previous studies (Hemminki et al., 2011; Hemminki, Sundquist, et al., 2016; Samadder et al., 2016). We found that only kidney cancer appeared to be associated with lower CUP risk, however only three CUP cases were available for analysis, so it is likely to be a chance finding. Previous associations observed between CUP and family history may possibly be explained by the general tendency for cancers of varying and discordant sites to cluster within families, rather than the family history itself directly increasing CUP risk. The most consistent association we observed was a marginally increased CUP risk in those with a sister with any cancer compared to those without sisters with cancer. The risk was moderately increased in age-sex adjusted models and multivariable adjusted models, and it remained statistically significant when restricting to histologically confirmed CUP cases. This finding may suggest that CUP is associated with cancers that occur in females, such as breast, uterine and ovarian cancer. However, we observed no associations between CUP and these cancers, so it is unlikely that the association seen in sisters is explained by female specific cancers. Instead, it is more likely that the association can be explained by sex specific excesses at other cancer sites such as lung cancer.

Table 3 Hazard ratios and 95% confidence intervals for Cancer of Unknown Primary risk in participants with family history of cancer in specific relatives and specific sites in those relatives in the Netherlands Cohort Study

		Subcohort members	Cancer of Unknown Primary cases					
			Cases	Age and sex-adjusted [1]			Multivariable adjusted [2]	
Cancer site in family member		Person time at risk (years)	n	HR	95% CI		HR	95% CI
Breast								
	No	66 211	871	1	Reference		1	Reference
	Yes	6 131	92	1.13	(0.88-	1.44)	1.15	(0.90- 1.48)
Ovarian								
	No	72 258	960	1	Reference		1	Reference
	Yes	84	3	2.50	(0.57-	10.86)	2.01	(0.36- 11.38)
Uterine								
	No	70 497	938	1	Reference		1	Reference
	Yes	1 845	25	1.00	(0.64-	1.58)	1.05	(0.67- 1.67)
Bowel								
	No	68 233	899	1	Reference		1	Reference
	Yes	4 109	64	1.15	(0.86-	1.54)	1.18	(0.88- 1.59)
Stomach								
	No	67 559	889	1	Reference		1	Reference
	Yes	4 783	74	1.11	(0.84-	1.46)	1.14	(0.87- 1.51)
Lung								
	No	65 336	880	1	Reference		1	Reference
	Yes	7 006	83	0.90	(0.70-	1.16)	0.89	(0.69- 1.15)
Kidney								
	No	71 517	960	1	Reference		1	Reference
	Yes	825	3	0.29	(0.09-	0.95)	0.27	(0.08- 0.90)
Prostate								
	No	70 687	938	1	Reference		1	Reference
	Yes	1 655	25	1.13	(0.72-	1.78)	1.20	(0.76- 1.89)

Cancer site in family member		Subcohort members	Cancer of Unknown Primary cases					
		Person time at risk (years)	Cases	Age and sex-adjusted [1]			Multivariable adjusted [2]	
			n	HR	95% CI		HR	95% CI
Bladder								
	No	71 555	951	1	Reference		1	Reference
	Yes	787	12	1.14	(0.60-	2.17)	1.17	(0.61- 2.26)
Pancreas								
	No	71 639	951	1	Reference		1	Reference
	Yes	703	12	1.45	(0.76-	2.75)	1.38	(0.72- 2.66)
Head and neck								
	No	71 025	948	1	Reference		1	Reference
	Yes	1 317	15	0.83	(0.47-	1.47)	0.81	(0.45- 1.44)
Leukaemia								
	No	70 260	937	1	Reference		1	Reference
	Yes	2 082	26	0.98	(0.64-	1.52)	0.99	(0.64- 1.55)
Lymphoma								
	No	71 706	958	1	Reference		1	Reference
	Yes	636	5	0.59	(0.23-	1.52)	0.59	(0.23- 1.55)

Notes

[1] Analyses were adjusted for age at baseline (years) and sex, and age as a time-varying covariate.

[2] Multivariable analyses were adjusted for age at baseline (years), sex, alcohol consumption (g ethanol/day), current cigarette smoking, cigarette smoking frequency (N/day; continuous; centered), cigarette smoking duration (years; continuous; centered), and cigarette smoking status (never/ever), cigarette smoking duration (continuous; centered) as time-varying covariates.

The strengths of this study lie in its prospective design, large cohort size, and high number of CUP cases available for analyses (compared to previous studies). Moreover, the data obtained from the NCR ensured that CUP cases were uniformly recorded and coded by trained registry clerks. Our study offers one particular advantage over previous studies, in that we were able to adjust for multiple confounders when estimating CUP risk. Addressing these confounders is essential as these lifestyle related factors (such as smoking and alcohol consumption) may modulate CUP risk, which could explain the marked associations in the Swedish studies. However, it should be noted that their methods to establish a participant's family history of cancer status may be more valid than those used in our study, as they were able to use the same registry to identify CUP cases and cancer in the family (Hemminki et al., 2011; Hemminki, Sundquist, et al., 2016). Furthermore, the use of a one-time measurement of presence of family history of cancer at baseline, may lead to non-differential misclassification of the participant's exposure status; participants may not report a family history of cancer at baseline, yet they may have family members diagnosed with cancer during the course of the follow-up. This misclassification may be augmented by the use of a questionnaire, relying on recall and close family ties, especially without verification of documented diagnoses in family members as in this study. This problem is likely to be increased if participants were asked to recall more specific details regarding the cancer site; it is easier for participants to recall whether their family members had cancer or not, rather than recall whether it was ovarian cancer or metastatic cancer (Schrijvers, Stronks, van de Mheen, Coebergh, & Mackenbach, 1994).

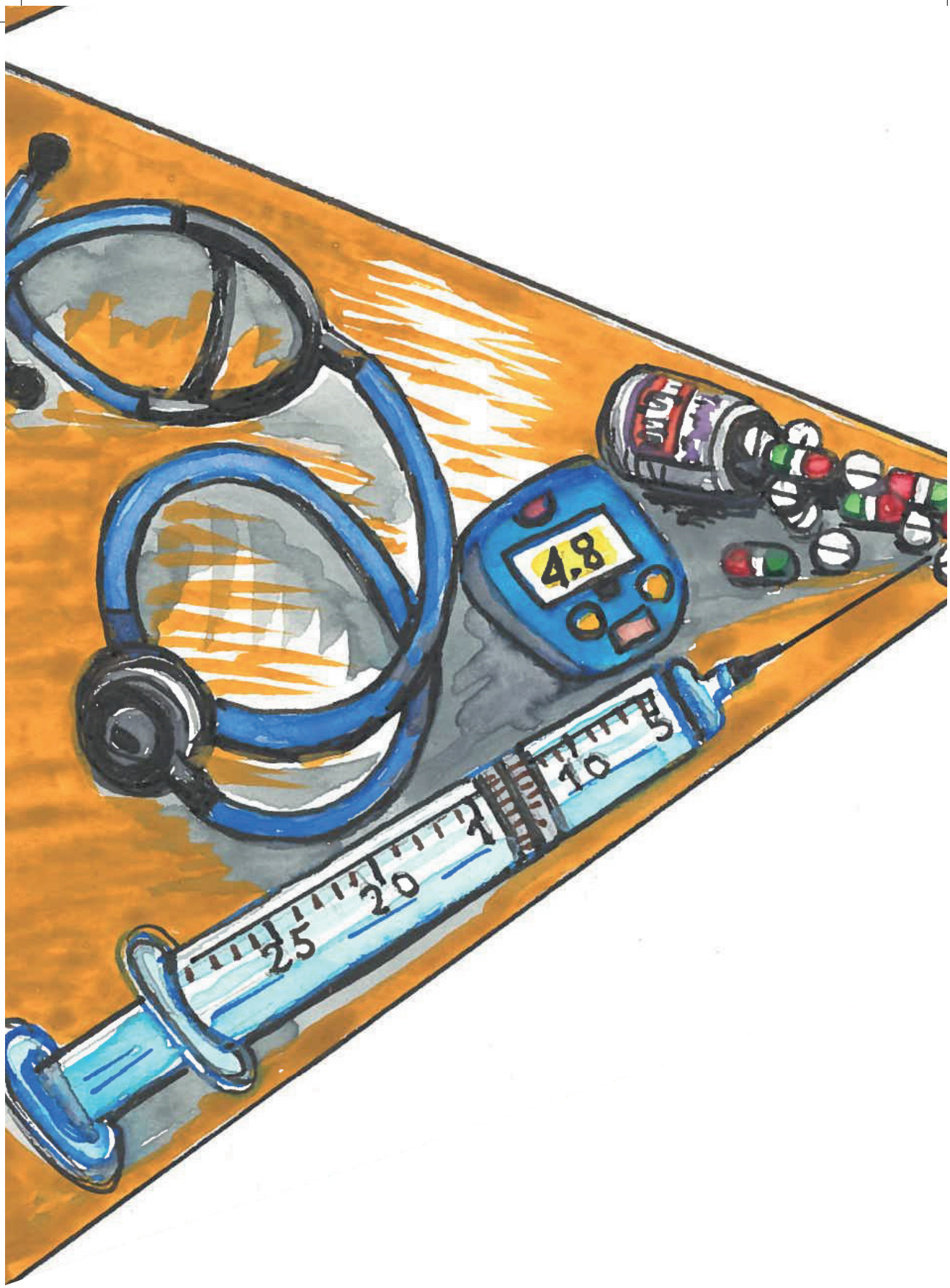
It has previously been highlighted that some familial cancers have a tendency for a younger age of diagnosis, and it is possible that any familial mechanism in CUP may present a similar pattern (Hemminki et al., 2011). This finding may explain the slightly higher estimates we observed between a family history of cancer and age. With our dataset being composed of those between the ages of 55-69, while CUP can occur at younger ages, it is possible that CUP cases where family history played a more prominent role might not have been available in our study population. Therefore, it remains highly plausible that this unavailability markedly reduced associations between family history and CUP in the NLCS.

To the best of our knowledge, this study is the first to have examined whether the presence of cancer in a person's family history affects their CUP risk. We thus provide new evidence to help uncover the role of familial aspects in CUP development. Within this cohort, having a family history of cancer is not an independent risk factor of CUP. In light of our findings, we suggest caution be employed when attempting to draw conclusions as to whether a family history of cancer increases CUP risk.

References

1. Barlow, W. E., Ichikawa, L., Rosner, D., & Izumi, S. J. J. o. c. e. (1999). Analysis of case-cohort designs. *52*(12), 1165-1172.
2. Comprehensive Cancer Centre the Netherlands. (2020). Incidentie Primaire Tumor Onbekend. Retrieved from https://iknl.nl/kankersoorten/primaire-tumor-onbekend/registratie/incidentie].
3. Hemminki, K., Bevier, M., Sundquist, J., & Hemminki, A. (2012). Cancer of unknown primary (CUP): does cause of death and family history implicate hidden phenotypically changed primaries? *Ann Oncol, 23*(10), 2720-2724. doi:10.1093/annonc/mds063
4. Hemminki, K., Chen, B., Melander, O., Manjer, J., Hallmans, G., & Hemminki, A. (2015). Smoking and body mass index as risk factors for subtypes of cancer of unknown primary. *Int J Cancer, 136*(1), 246-247. doi:10.1002/ijc.28969
5. Hemminki, K., Försti, A., Sundquist, K., & Li, X. (2016). Cancer of unknown primary is associated with diabetes. *Eur J Cancer Prev., 25*(3), 246-251.
6. Hemminki, K., Ji, J., Sundquist, J., & Shu, X. (2011). Familial risks in cancer of unknown primary: tracking the primary sites. *J Clin Oncol, 29*(4), 435-440. doi:10.1200/JCO.2010.31.5614
7. Hemminki, K., Sundquist, K., Sundquist, J., Hemminki, A., & Ji, J. (2016). Location of metastases in cancer of unknown primary are not random and signal familial clustering. *Scientific reports, 6*(1), 1-8.
8. Hemminki, K., Sundquist, K., Sundquist, J., & Ji, J. (2015). Risk of cancer of unknown primary after hospitalization for autoimmune diseases. *Int J Cancer, 137*(12), 2885-2895. doi:10.1002/ijc.29657
9. Hermans, K. E., van den Brandt, P. A., Loef, C., Jansen, R. L., & Schouten, L. J. (2021). Alcohol consumption, cigarette smoking and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study. *Int J Cancer*.
10. Kaaks, R., Sookthai, D., Hemminki, K., Kramer, A., Boeing, H., Wirfalt, E., . . . Canzian, F. (2014). Risk factors for cancers of unknown primary site: Results from the prospective EPIC cohort. *Int J Cancer, 135*(10), 2475-2481. doi:10.1002/ijc.28874
11. Lin, D. Y., & Wei, L.-J. (1989). The robust inference for the Cox proportional hazards model. *J Am Stat Assoc, 84*(408), 1074-1078.
12. Luke, C., Koczwara, B., Karapetis, C., Pittman, K., Price, T., Kotasek, D., . . . Roder, D. (2008). Exploring the epidemiological characteristics of cancers of unknown primary site in an Australian population: implications for research and clinical care. *Aust N Z J Public Health, 32*(4), 383-389. doi:10.1111/j.1753-6405.2008.00260.x
13. National Institute for Health and Care Excellence. (2010). *Metastatic Malignant Disease of Unknown Primary Origin. Diagnosis and Management of Metastatic Malignant Disease of Unknown Primary Origin*: NICE.
14. Pavlidis, N., Rassy, E., & Smith-Gagen, J. (2020). Cancer of unknown primary: Incidence rates, risk factors and survival among adolescents and young adults. *Int J Cancer, 146*(6), 1490-1498.
15. Rassy, E., Assi, T., & Pavlidis, N. (2020). Exploring the biological hallmarks of cancer of unknown primary: where do we stand today? *Br J Cancer*, 1-9.
16. Rassy, E., & Pavlidis, N. (2019). The currently declining incidence of cancer of unknown primary. *J Cancer Epidemiol, 61*, 139-141.

17. Samadder, N. J., Smith, K. R., Hanson, H., Pimentel, R., Wong, J., Boucher, K., . . . Burt, R. W. (2016). Familial risk in patients with carcinoma of unknown primary. *JAMA Oncology, 2*(3), 340-346.
18. Schrijvers, C. T., Stronks, K., van de Mheen, D. H., Coebergh, J.-W. W., & Mackenbach, J. P. (1994). Validation of cancer prevalence data from a postal survey by comparison with cancer registry records. *Am J Epidemiol, 139*(4), 408-414.
19. Schroten-Loef, C., Verhoeven, R., de Hingh, I., van de Wouw, A., van Laarhoven, H., & Lemmens, V. (2018). Unknown primary carcinoma in the Netherlands: decrease in incidence and survival times remain poor between 2000 and 2012. *Eur. J. Cancer, 101*, 77-86.
20. Vajdic, C., Perez-Concha, O., Dobbins, T., Ward, R. L., Schaffer, A. L., van Leeuwen, M., . . . Pearson, S.-A. (2019). Demographic, social and lifestyle risk factors for cancer registry-notified cancer of unknown primary site (CUP). *Cancer Epidemiol., 60*, 156-161.
21. Van den Brandt, P. A., Goldbohm, R. A., van 't Veer, P., Volovics, A., Hermus, R. J., & Sturmans, F. (1990). A large-scale prospective cohort study on diet and cancer in The Netherlands. *J Clin Epidemiol, 43*(3), 285-295.
22. Van den Brandt, P. A., Schouten, L. J., Goldbohm, R. A., Dorant, E., & Hunen, P. M. (1990). Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. *Int J Epidemiol, 19*(3), 553-558.
23. Zeegers, M. P., Schouten, L. J., Goldbohm, R. A., & van den Brandt, P. A. (2008). A compendium of familial relative risks of cancer among first degree relatives: A population-based study. *Int J Cancer, 123*(7), 1664-1673.


4.8
25
20
15
10
5

CHAPTER 7

TYPE 2 DIABETES MELLITUS AND CANCER OF UNKNOWN PRIMARY RISK

Karlijn E.P.E. Hermans, Sharmi Haque, Piet A. van den Brandt, Caroline Loef, Rob L. H. Jansen, and Leo J. Schouten

Eur. J. Cancer Prev. 2022

Abstract

Objectives: Cancer of unknown primary (CUP) is a metastatic malignancy with an unidentifiable primary tumour origin. Previous studies suggest that Type 2 diabetes mellitus (T2DM) and CUP risk are associated. This study examines the association in greater depth by investigating T2DM status, T2DM duration, T2DM age at diagnosis, and medication that was being used in relation to CUP.

Methods: Data was utilised from the Netherlands Cohort Study, a prospective cohort that includes 120,852 participants aged 55-69 years at baseline in 1986. All participants completed a self-administered questionnaire on cancer risk factors. CUP cases were identified through record linkage with the Netherlands Cancer Registry and Dutch Pathology Registry. After 20.3 years of follow-up, 963 incident CUP cases and 4,288 subcohort members were available for case-cohort analyses. Proportional hazards models were employed to estimate multivariable adjusted hazard ratios (HRs).

Results: Overall, we observed a non-significant positive association between T2DM status and CUP risk (HR: 1.35, 95% CI: 0.92-1.99), which increased in women after stratification for sex (HR: 1.55, 95% CI: 0.90-2.64). For participants who were aged <50 years at diagnosis of T2DM, a statistically significant positive association was found in relation to CUP (HR: 2.42, 95% CI: 1.26-4.65), compared to participants without T2DM.

Conclusions: Our findings indicate that there is a non-significant positive association between T2DM and CUP risk, and that the association became stronger in women in stratified analyses.

Introduction

Cancer of Unknown Primary (CUP) denotes a heterogenous group of metastatic tumours where the identification of the primary tumour is unknown at diagnosis (Loffler and Kramer, 2016). Globally, CUP incidence rates have been declining and currently reaches 1-2% of all incident cancer diagnoses (Rassy and Pavlidis, 2019). In the Netherlands, approximately 1,300 patients have been diagnosed with CUP in 2018 (Schroten-Loef et al., 2018,Comprehensive Cancer Centre the Netherlands, 2020,Meijer et al., 2020). As yet, little research has been conducted into its aetiology. CUP risk factors that have been identified in previous epidemiological studies include smoking (Kaaks et al., 2014,Hemminki et al., 2015,Vajdic et al., 2019a,Hermans et al., 2021a), alcohol consumption (Hermans et al., 2021a), and meat consumption (Vajdic et al., 2019a,Hermans et al., 2021b).

In the current study, we are particularly interested to explore the relationship between Type 2 diabetes mellitus (T2DM) and CUP risk, as T2DM has been associated with increased risk of various types of cancer (Giovannucci et al., 2010). T2DM is characterised by metabolic disorders that are denoted by hyperglycaemia, and accounts for approximately 90% of all DM (Giovannucci et al., 2010,Nolan et al., 2011,Saeedi et al., 2019). Worldwide, it was estimated that 463 million people had DM (Type 1 and 2 combined) in 2019, and numbers are expected to increase rapidly (Saeedi et al., 2019). To the best of our knowledge, only two studies have investigated the association between T2DM and CUP risk (Hemminki et al., 2016,Vajdic et al., 2019b). A Swedish study found that T2DM was associated with increased CUP risk compared to participants with no diabetes (Hemminki et al., 2016). In an Australian prospective cohort study, patients with DM were also found to be at an increased CUP risk (Vajdic et al., 2019b). We aimed to further examine the association between T2DM and CUP risk by assessing T2DM duration, age at diagnosis, and medication that was being used, as well as multiplicative interactions between smoking, BMI, physical activity, and alcohol consumption in relation to CUP risk.

Methods

Design and study population

The Netherlands Cohort Study (NLCS) commenced in 1986 and included 120,852 participants between 55-69 years at baseline. Participants originated from 204 Dutch municipalities. Participants became part of the cohort after returning their completed questionnaire ([Van den Brandt et al., 1990b]). Data processing and analysis was conducted through a case-cohort approach where incident CUP cases were obtained from the full cohort through record linkage with the Netherlands Cancer Registry (NCR) and the Dutch Pathology Registry (PALGA). The number of person-years at risk for the full cohort was estimated from a subcohort of 5,000 participants which were randomly sampled from the full cohort at baseline.

Outcome measure

CUP is defined as a metastasised epithelial malignancy with no identifiable primary tumour origin after cytological and/or histological verification during a patient's lifetime. This CUP definition only includes epithelial malignancies (ICD-O-3: M-8000 - M-8570), which excludes sarcoma, lymphoma, mesothelioma, and melanoma.

Follow-up

Cancer follow-up was determined through annual record linkage of the full cohort with the Netherlands Cancer Registry (NCR) and the Dutch Pathology Registry (PALGA) to identify CUP cases within the NLCS ([Van den Brandt et al., 1990a]). Follow-up duration was 20.3 years (17 September 1986 until 31 December 2006, which resulted in 1,353 potential CUP cases and 4774 subcohort members who did not report a history of cancer (except for skin cancer) at baseline. Participants were excluded from the analyses if there was missing data on selected confounders (390 CUP cases and 486 subcohort members). Consequently, data were available for 963 incident CUP cases with microscopical confirmation/and epithelial histology and 4,288 subcohort members.

Questionnaire data

All participants filled out a self-administered questionnaire on dietary habits and other risk factors for cancer at baseline. With respect to DM, the questionnaire addressed the following questions: 'Has a physician ever diagnosed you with diabetes mellitus and what was your age at that time?' Participants could select the corresponding age category ranging from 'younger than 30 years', followed by 5-year age categories ranging from '30 to 34 years' up to '65 to 69 years'. Based on previous epidemiological evidence, we determined that if participants indicated to have been diagnosed with DM after the age of 30 years, they were classified as having T2DM ([de Kort et al., 2016]). Participants with probable Type 1 diabetes mellitus (T1DM) were excluded from analyses (8 CUP cases and 10 subcohort members). Diabetes duration was calculated by subtracting the age at diagnosis of DM from the age at baseline. Participants were also asked to indicate 'What medication they used longer than six months, for what condition(s) and in what period(s)?'. They could fill in the name of the medication and for what condition the medicine was used in their respective time period(s). The medication was classified according to the Anatomical Therapeutic Chemical (ATC) from the World Health Organization Collaborative Centre for Drug Statistical Methodology. Anti-diabetic medication was categorised into drugs based on "insulin and analogues" and "drugs lowering the blood glucose level (excluding insulin)". The questionnaire also included detailed questions on smoking, BMI, physical activity and alcohol consumption. Smoking was measured through questions related to baseline smoking status, age at first exposure and last exposure to smoking after cessation. Smoking frequency and smoking duration (excluding cessation periods) for cigarette, cigar and pipe smokers. Participants that were considered themselves to be non-smokers were denoted as never smokers. To avoid collinearity problems, smoking frequency and smoking duration were centered as proposed by Leffondré et al. ([Leffondre et al., 2002]). Height in centimetres (cm) and weight in kilogram (kg) were determined at baseline and permitted the calculation of BMI at baseline squared (kg/m^2). Physical activity was examined by questions encircling recreational physical activity, and the physical activity involved in commuting to and from work (e.g., walking and cycling). The reported times from the participants were generated into a total non-occupational physical activity value. Habitual alcohol consumption was measured over the year by addressing questions on beer, red wine, white wine, sherry and other fortified wine, liqueur and liquors at baseline. Frequency of alcohol consumptions was measured in pre-determined

ranges from 'never' to '6-7 times per week'. Information was requested through questions relating to the number of glasses consumed on a daily basis. Abstainers were considered as participants that indicated that they consumed 'less than once per month' or 'never'. Mean daily alcohol consumption was calculated by using the computerised Dutch Food composition table ([NEVO-table, 1986]).

Ethical information

Participants consented to be included in the cohort and follow-up by returning their completed questionnaires. The institutional review boards of the Netherlands Organization for Applied Scientific Research TNO (Zeist) and Maastricht University (Maastricht) approved the execution of the NLCS and the informed consent procedure. The study complies with the medical ethical standards of the Declaration of Helsinki.

Statistical information

CUP cases were obtained from the full cohort. For the subcohort, the person time at risk was calculated from baseline (17 September 1986) until CUP diagnosis, death, emigration, loss to follow-up or end of follow-up (31 December 2006), whichever occurred first. General characteristics were presented for both subcohort members and CUP cases with frequencies and percentages for categorical variables and means including standard deviations for continuous variables.

T2DM was assessed based on status (yes/no), duration (0-10 years, and >10 years), age at diagnosis (<50 years, and ≥50 years) and use of medication (no medication, insulin and analogues treatment, or the use of blood glucose lowering drugs (excluding insulin). Predefined confounders included age at baseline (years; continuous); sex (male/female); alcohol consumption (ethanol intake in grams per day); cigarette smoking status (never/ever); cigarette smoking frequency (number of cigarettes smoked per day; centered); and cigarette smoking duration (number of years smoking; centered). Potential confounders included body mass index (BMI) at baseline (kg/m^2); non-occupational physical activity (<30 min/day, 30-60 min/day, 60-90 min/day and >90 min/day); socio-economic status (highest level of education); and history of cancer in a first degree relative (yes/no). For the final model, we have included the predefined confounders alone, as none of the potential confounders altered the HRs by >10%.

Wald tests and cross-product terms were used to evaluate potential multiplicative interaction with respect to 1) sex, 2) cigarette smoking status, 3) BMI at baseline, 4) physical activity, and 5) alcohol consumption in relation to CUP risk. Age- and sex-adjusted, and multivariable-adjusted HRs with 95% confidence intervals were estimated using Cox proportional hazards models. Standard errors were calculated using the robust Huber-White sandwich estimator to account for additional variance introduced by sampling from the full cohort ([Barlow, 1994]). The proportional hazards assumption was tested using the scaled Schoenfeld residuals ([Lin and Wei, 1989]), and by visual inspection of log-minus-log survival curves. If the assumption had been violated, a time-varying covariate for that variable was added to the model where appropriate. Ordinal exposure variables were fitted as continuous variables in trend analyses. All analyses were conducted using Stata version 15. *P*-values were considered statistically significant if $p < 0.05$.

A sensitivity analysis was performed by restricting the analysis to histologically verified CUP cases alone. For histologically verified CUP patients, it is more likely that they underwent extensive diagnostic investigation(s) to rule out the primary tumour origin. For patients who received cytological verification alone, other factors may have played a role in the decision to refrain from further diagnostic investigation such as age, comorbidities, performance status, localisation of metastasis, or the patient's decision.

Results

Analyses were conducted using 963 incident CUP cases and 4,288 subcohort members. General characteristics of subcohort members and CUP cases display that there were more men diagnosed with CUP 62.6% than were women (see Table 1). T2DM status was slightly more prevalent in women than in men in both the CUP cases and the subcohort (5.0% and 3.8%, & 3.5% and 3.4%, respectively).

We observed a multiplicative interaction between T2DM duration and sex ($P_{interaction}$ = 0.03), whereas no multiplicative interactions were observed between T2DM status and sex, cigarette smoking, BMI, physical activity, or alcohol consumption in relation to CUP risk ($P_{interaction}$ = 0.35, 0.37, 0.08, 0.53 and 0.34), respectively. We have, therefore, conducted sex-stratified analyses based on the finding of the multiplicative interaction between T2DM duration and sex (see Table 2 and Table 3). Overall, there appears to be a slightly increased risk between T2DM and CUP (HR:

1.35, 95% CI: 0.92-1.99) compared to participants who did not have T2DM, although it was not statistically significant. After stratification for sex, the association increased in women (HR: 1.55, 95% CI: 0.90-2.64), whereas it attenuated in men (HR: 1.19, 95% CI: 0.69-2.04). Furthermore, it appeared that participants who had T2DM for >10 years had a higher CUP risk (HR: 1.49, 95% CI: 0.82-2.70) compared to participants who did not have T2DM, but again this finding was not statistically significant. Participants who were aged <50 years at diagnosis of T2DM had an increased CUP risk that was statistically significant (HR: 2.42, 95% CI: 1.26-4.65, P_{trend} =0.03), while no association was found for participants aged ≥50 years at diagnosis. The association appeared to be increased in women, whom were diagnosed with T2DM (<50 years), as they had a significant increased CUP risk (HR: 2.72, 95% CI: 1.13-6.55) compared to women without T2DM. While in men, a non-significant increased CUP risk (HR: 2.15, 95% CI: 0.85-5.42) was seen. Furthermore, we observed a non-significant association in participants with T2DM who did not use medication in relation to CUP risk (HR: 1.42, 95% CI: 0.85-2.38) compared to not having T2DM. Participants who have T2DM and use blood glucose lowering drugs (excluding insulin) had a slightly increased CUP risk (HR: 1.24, 95% CI: 0.66-2.32), although not statistically significant, while we observed no association between participants who have T2DM and use insulin and analogues treatment compared to participants who did not have T2DM. It should be acknowledged that there were only few participants who had T2DM and that used blood glucose lowering drugs (excluding insulin) (n=14), as well as participants who had T2DM and used insulin and analogues treatment (n=3).

The sensitivity analysis results were confined to histologically verified CUP which comprised data on 693 CUP cases, the findings reflected those of the complete multivariable analysis (data not shown).

Table 1 Characteristics of Cancer of Unknown Primary cases and subcohort members in the Netherlands Cohort Study

	Subcohort members (n=4288)		Cancer of Unknown Primary cases (n=963)	
Characteristic	**n**	**(%)**	**n**	**(%)**
Age at baseline (years)				
55-59	1664	38.8	288	29.9
60-64	1461	34.1	372	38.6
65-69	1163	27.1	303	31.5
Sex				
Men	2110	49.2	603	62.6
Women	2178	50.8	360	37.4
Type 2 diabetes status	153	3.6	39	4.1
Men	71	3.4	21	3.5
Women	82	3.8	18	5.0
Ethanol intake (grams/day) [1]				
Abstainers	1024	23.9	186	19.3
<5	1228	28.6	247	25.7
5-<15	979	22.8	217	22.5
15-<30	672	15.7	153	15.9
≥30	385	9.0	160	16.6
Cigarette smoking status				
Never smokers	1584	36.9	265	27.5
Ex smokers	1521	35.5	334	34.7
Current smokers	1183	27.6	364	37.8
Frequency of cigarette smoking (N/day), mean (SD) [2]	15.7 (10.1)		17.8 (10.4)	

	Subcohort members (n=4288)		Cancer of Unknown Primary cases (n=963)	
Characteristic	**n**	**(%)**	**n**	**(%)**
Duration of cigarette smoking (years), mean (SD) [2]	31.9 (12.1)		35.6 (11.6)	
Body Mass Index at baseline (kg/m²), mean (SD)	25.0 (3.1)		25.0 (3.0)	
Non-occupational physical activity (min/day)				
≤30	908	21.5	204	21.5
>30-60	1318	31.2	291	30.6
>60-90	879	20.8	170	17.9
>90	1122	26.5	285	30.0
Level of education (years of education)				
Primary	1257	29.5	271	28.5
Lower vocational	937	22.0	204	21.4
Secondary and medium vocational	1483	34.8	341	35.8
University and higher vocational	590	13.8	136	14.3
First grade family history of cancer				
Yes	1945	45.4	459	47.7

Notes

[1] In consumers only

[2] In users only

Table 2 Hazard ratios and 95% confidence intervals for diabetes and Cancer of Unknown Primary risk in the Netherlands Cohort Study

	Subcohort members (n=4288)	Cancer of Unknown Primary cases (n=963)				
	Person time at risk (years)	Cases	Age- and sex-adjusted [1]		Multivariable adjusted [2]	
		n	HR	95% CI	HR	95% CI
T2DM status						
No	70 144	924	1	Reference	1	Reference
Yes	2 198	39	1.31	(0.90- 1.91)	1.35	(0.92- 1.99)
p for interaction with sex					0.35	
Duration T2DM						
No T2DM	70 144	924	1	Reference	1	Reference
T2DM (0-10 years)	1 357	23	1.27	(0.78- 2.06)	1.27	(0.77- 2.08)
T2DM (>10 years)	841	16	1.37	(0.76- 2.46)	1.49	(0.82- 2.70)
p for interaction with sex					0.03	
p for trend [3]			0.17		0.11	
Age at diagnosis T2DM						
No T2DM	70 144	924	1	Reference	1	Reference
T2DM (diagnosed ≥50 years)	1 644	24	1.04	(0.65- 1.66)	1.06	(0.66- 1.70)
T2DM (diagnosed <50 years)	554	15	2.20	(1.16- 4.19)	2.42	(1.26- 4.65)
p for interaction with sex					0.57	
p for trend [3]			0.05		0.03	
T2DM medication						
No T2DM	70 079	924	1	Reference	1	Reference
T2DM no medication	1 148	22	1.40	(0.85- 2.33)	1.42	(0.85- 2.38)
T2DM with blood glucose lowering drugs (excluding insulin)	845	14	1.19	(0.65- 2.20)	1.24	(0.66- 2.32)
T2DM with insulin and analogues treatment	270	3	0.91	(0.26- 3.15)	1.03	(0.29- 3.63)
p for interaction with sex					0.50	

Notes

[1] Analyses were adjusted for age at baseline (years) and sex.

[2] Analyses were adjusted for age at baseline (years), sex, alcohol consumption, cigarette smoking status (never/ever), cigarette smoking frequency (continuous; centered), and cigarette smoking duration (continuous; centered). Cigarette smoking duration was additionally included as a time-varying covariate.

[3] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

Table 3 Hazard ratios and 95% confidence intervals for diabetes and Cancer of Unknown Primary risk in the Netherlands Cohort Study, stratified for sex

	Men only							Women only						
	Sub cohort members (n=2110)	Cancer of Unknown Primary cases (n=603)						Sub cohort members (n=2178)	Cancer of Unknown Primary cases (n=360)					
	Person time at risk (years)	# n	Age- and sex-adjusted [1] HR	95% CI		Multivariable adjusted [2] HR	95% CI	Person time at risk (years)	# n	Age- and sex-adjusted [1] HR	95% CI		Multivariable adjusted [2] HR	95% CI
T2DM status														
No	32 640	582	1	Reference		1	Reference	37 504	342	1	Reference		1	Reference
Yes	888	21	1.20	(0.71-	2.03)	1.19	(0.69- 2.04)	1 310	18	1.46	(0.85-	2.48)	1.55	(0.90- 2.64)
Duration T2DM														
No T2DM	32 640	582	1	Reference		1	Reference	37 504	342	1	Reference		1	Reference
T2DM (0-10 years)	523	13	1.29	(0.66-	2.51)	1.19	(0.60- 2.37)	834	10	1.24	(0.62-	2.50)	1.33	(0.66- 2.67)
T2DM (>10 years)	365	8	1.07	(0.47-	2.47)	1.18	(0.51- 2.75)	476	8	1.85	(0.83-	4.13)	1.95	(0.88- 4.36)
p for trend [3]			0.61			0.56				0.11			0.08	
Age at diagnosis T2DM														
No T2DM	32 640	582	1	Reference		1	Reference	37 504	342	1	Reference		1	Reference
T2DM (diagnosed ≥50 years)	662	13	0.98	(0.51-	1.87)	0.93	(0.48- 1.80)	982	11	1.13	(0.58-	2.19)	1.21	(0.62- 2.35)
T2DM (diagnosed <50 years)	226	8	1.88	(0.77-	4.61)	2.15	(0.85- 5.42)	328	7	2.69	(1.12-	6.49)	2.72	(1.13- 6.55)
p for trend [3]			0.31			0.28				0.06			0.04	
T2DM medication														
No T2DM	32 605	582	1	Reference		1	Reference	37 474	342	1	Reference		1	Reference
T2DM no medication	448	12	1.29	(0.63-	2.60)	1.25	(0.61- 2.58)	701	10	1.55	(0.76-	3.17)	1.63	(0.80- 3.33)
T2DM with blood glucose lowering drugs (excluding insulin)	385	7	0.96	(0.40-	2.29)	0.98	(0.40- 2.39)	460	7	1.55	(0.67-	3.62)	1.63	(0.70- 3.82)
T2DM with insulin and analogues treatment	90	2	1.29	(0.26-	6.47)	1.49	(0.28- 7.88)	180	1	0.57	(0.07-	4.52)	0.62	(0.08- 4.97)

Notes

[1] Analyses were adjusted for age at baseline (years) and sex.

[2] Analyses were adjusted for age at baseline (years), sex, alcohol consumption, cigarette smoking status (never/ever), cigarette smoking frequency (continuous; centered), and cigarette smoking duration (continuous; centered). Cigarette smoking duration was additionally included as a time-varying covariate.

[3] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

Discussion

In this prospective study, we examined the association between T2DM and CUP risk. Overall, we observed a non-significant positive association between T2DM status and CUP. After stratification for sex, the association between T2DM status and CUP risk became stronger in women. Participants who had T2DM for >10 years appeared to have a higher CUP risk, although the association was not statistically significant. For participants who were aged <50 years at diagnosis of T2DM, a statistically significant positive association was found in relation to CUP. We observed a multiplicative interaction between T2DM duration and sex, whereas no interactions were found between T2DM status and sex, cigarette smoking, BMI, physical activity, or alcohol consumption in relation to CUP.

To the best of our knowledge, only two studies previously investigated the association between diabetes and CUP. One study, a Swedish study, investigated patients with T2DM (51,929 cases with insulin treatment, 126,515 cases without insulin treatment) who were identified from the national healthcare registers, and linked to the Swedish Cancer Registry. Its authors computed standardized incidence ratios (SIRs), and reported associations between T2DM with insulin treatment (SIR: 1.38, 95% CI: 1.12-1.67) and T2DM without insulin treatment (SIR: 1.78, 95% CI: 1.58-2.00) in relation to CUP risk compared to participants with no diabetes (Hemminki et al., 2016). An Australian cohort study, compared 327 incident CUP cases to two sets of randomly selected controls (3:1). In this Australian study, diabetes was measured as a self-reported health condition. Its authors found a statistically significantly association between diabetes and CUP risk in both control groups; metastatic cancer known primary controls (odds ratio [OR]: 1.58, 95% CI: 1.11-2.26) and general cohort population controls (OR: 2.36, 95% CI: 1.54-3.62) compared to participants with no diabetes (Vajdic et al., 2019b). Although not statistically significant, we also observed a slightly increased CUP risk, and a stronger association with T2DM without medication use in the NLCS. Hence, our point estimates appear to be in agreement with the Swedish and Australian studies (Hemminki et al., 2016,Vajdic et al., 2019b). In general, patients with T2DM have an impaired immune system, which can result in improper immune response (Berbudi et al., 2020). Therefore, patients may have an increased vulnerability when it comes to cancer development overall. For CUP patients specifically, it may be proposed that the immune system was able to supress the primary tumour, while the metastasis escaped suppression due to improper immune response (Pavlidis and Pentheroudakis, 2012,Hemminki et al., 2016,Loffler and Kramer, 2016).

Within the NLCS, we explored the association between T2DM and CUP risk by multiplicative interactions as well as T2DM duration, and age at diagnosis of T2DM. We found a multiplicative interaction between T2DM duration and sex, consequently, we have additionally examined sex-stratified associations with respect to T2DM. Overall, we observed that participants who had T2DM for >10 years had a higher non-significant CUP risk, compared to participants with no diabetes. This association became markedly stronger when restricting the analysis to women alone. Moreover, we found a statistically significant positive association for participants who were aged <50 years at diagnosis of T2DM in relation to CUP risk. The observed association became stronger and statistically significant in women alone. In an extensive review on diabetes and all-site cancer events, it was reported that women had a 6% higher excess-risk than men ([Ohkuma et al., 2018]). Its authors also gave possible explanations for differences in this excess-risk; being that women with diabetes may have a poorer glycaemic control compared to men with diabetes, and hyperinsulinemia may be longer in women than in men as the average duration of impaired glucose tolerance was reported to be >2 years longer in women compared to men, which may have played an important role in the prediabetic period ([Bertram and Vos, 2010,Ohkuma et al., 2018]). Based on these explanations, our findings with respect to the sex-differences between T2DM and CUP risk, seem to be comparable to those of overall T2DM and cancer risk. It should, however, be acknowledged that this might be a chance finding due to the utilisation of multiple comparisons in categories with fewer participants as a result of stratification.

The strengths of this study are its prospective cohort design, relatively large cohort size of 120,852 participants, higher number of 963 incident CUP cases, lengthy follow-up time of 20.3 years, and wide availability of confounder data. Additionally, record linkage with the NCR and PALGA for cancer follow-up was complete for at least 96%, and vital status follow-up was complete for almost 100% after 20.3 years, thereby minimizing selection bias ([Goldbohm et al., 1994]). Besides, the registry clerks from the NCR applied uniform coding rules which enabled the use of a consistent disease definition based on pathology and clinical reports ([Van der Sanden et al., 1995]). There are a few limitations of our study, namely, the number of exposed cases (participants with T2DM) is relatively low and those cases derived from the baseline measurement alone (no follow-up). Also, data on diabetes status was obtained from the self-reported questionnaire by the participant, therefore, it is possible that we have missed a diabetes diagnosis during follow-up, or that a diagnosis was inaccurate

as a result from self-reporting. However, validation studies have indicated that self-reported diabetes status can be used as an accurate proxy for studying diabetes status ([Comino et al., 2013,Pastorino et al., 2015]). It is also important to acknowledge that the prevalence, awareness, and knowledge about DM, as well as its management, has increased and improved over the past 20 years ([Nazar et al., 2016]). At the start of the NLCS in 1986, patients with diabetes may have had undiagnosed (pre)diabetes which may underestimate our findings. The NLCS dataset does not enable us to check which diagnostic methods were applied for diagnosing CUP, however, if we restrict our analysis to histologically verified CUP cases alone, the results do not differ substantially from those of the overall multivariable analyses. Hence, we conclude that the findings of our multivariable analyses are representative of CUP cases with or without an extensive diagnostic work-up.

In conclusion, there appears to be a non-significant positive association between T2DM and CUP risk in the NLCS. After stratification for sex, the association between T2DM status and CUP risk became stronger in women.

References

1. Barlow WE (1994). Robust variance estimation for the case-cohort design. Biometrics:1064-1072.
2. Berbudi A, Rahmadika N, Tjahjadi AI, Ruslami R (2020). Type 2 Diabetes and its Impact on the Immune System. Current diabetes reviews 16(5):442-449. doi: 10.2174/1573399815666191024085838.
3. Bertram MY, Vos T (2010). Quantifying the duration of pre-diabetes. J Australian New Zealand journal of public health 34(3):311-314.
4. Comino EJ, Tran DT, Haas M, Flack J, Jalaludin B, Jorm L et al. (2013). Validating self-report of diabetes use by participants in the 45 and up study: a record linkage study. BMC Health Services Research 13(1):481. doi: 10.1186/1472-6963-13-481.
5. Comprehensive Cancer Centre the Netherlands. Incidentie Primaire Tumor Onbekend. Integraal kankercentrum Nederland; 2020.
6. de Kort S, Simons CC, van den Brandt PA, Goldbohm RA, Arts IC, de Bruine AP et al. (2016). Diabetes mellitus type 2 and subsite-specific colorectal cancer risk in men and women: results from the Netherlands Cohort Study on diet and cancer. European journal of gastroenterology & hepatology 28(8):896-903. doi: 10.1097/meg.0000000000000626.
7. Giovannucci E, Harlan DM, Archer MC, Bergenstal RM, Gapstur SM, Habel LA et al. (2010). Diabetes and cancer: a consensus report. Diabetes care 33(7):1674-1685. doi: 10.2337/dc10-0666.
8. Goldbohm RA, Van den Brandt PA, Dorant E (1994). Estimation of the coverage of Dutch municipalities by cancer registries and PALGA based on hospital discharge data. Tijdschr Soc Gezondheidsz 72:80-84.
9. Hemminki K, Försti A, Sundquist K, Li X (2016). Cancer of unknown primary is associated with diabetes. Eur J Cancer Prev 25(3):246-251.
10. Hemminki K, Chen B, Melander O, Manjer J, Hallmans G, Hemminki A (2015). Smoking and body mass index as risk factors for subtypes of cancer of unknown primary. Int J Cancer 136(1):246-247. doi: 10.1002/ijc.28969.
11. Hermans KE, van den Brandt PA, Loef C, Jansen RL, Schouten LJ (2021a). Alcohol consumption, cigarette smoking and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study. Int J Cancer.
12. Hermans KEPE, van den Brandt PA, Loef C, Jansen RLH, Schouten LJ (2021b). Meat consumption and cancer of unknown primary (CUP) risk: results from The Netherlands cohort study on diet and cancer. European Journal of Nutrition. doi: 10.1007/s00394-021-02600-5.
13. Kaaks R, Sookthai D, Hemminki K, Kramer A, Boeing H, Wirfalt E et al. (2014). Risk factors for cancers of unknown primary site: Results from the prospective EPIC cohort. Int J Cancer 135(10):2475-2481. doi: 10.1002/ijc.28874.
14. Leffondre K, Abrahamowicz M, Siemiatycki J, Rachet B (2002). Modeling smoking history: a comparison of different approaches. American journal of epidemiology 156(9):813-823. doi: 10.1093/aje/kwf122.
15. Lin DY, Wei L-J (1989). The robust inference for the Cox proportional hazards model. J Am Stat Assoc 84(408):1074-1078.

16. Loffler H, Kramer A (2016). Introduction. In: Cancer of Unknown Primary Loffler H, Kramer A (editors): Springer International Publishing Switzerland.
17. Meijer L, De Peuter R, Van der Zwan JM, Loef C. Primaire Tumor Onbekend. Integraal kankercentrum Nederland; 2020.
18. Nazar CM, Bojerenu MM, Safdar M, Marwat J (2016). Effectiveness of diabetes education and awareness of diabetes mellitus in combating diabetes in the United Kingdom; a literature review. Journal of nephropharmacology 5(2):110-115.
19. NEVO-table (1986). Dutch food composition table 1986–1987. Voorlichtingsbureau voor de voeding, The Hague, The Netherlands.
20. Nolan CJ, Damm P, Prentki M (2011). Type 2 diabetes across generations: from pathophysiology to prevention and management. The Lancet 378(9786):169-181. doi: https://doi.org/10.1016/S0140-6736(11)60614-4.
21. Ohkuma T, Peters SAE, Woodward M (2018). Sex differences in the association between diabetes and cancer: a systematic review and meta-analysis of 121 cohorts including 20 million individuals and one million events. Diabetologia 61(10):2140-2154. doi: 10.1007/s00125-018-4664-5.
22. Pastorino S, Richards M, Hardy R, Abington J, Wills A, Kuh D et al. (2015). Validation of self-reported diagnosis of diabetes in the 1946 British birth cohort. Primary care diabetes 9(5):397-400. doi: 10.1016/j.pcd.2014.05.003.
23. Pavlidis N, Pentheroudakis G (2012). Cancer of unknown primary site. Lancet 379(9824):1428-1435. doi: 10.1016/S0140-6736(11)61178-1.
24. Rassy E, Pavlidis N (2019). The currently declining incidence of cancer of unknown primary. J Cancer Epidemiol 61:139-141.
25. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N et al. (2019). Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas. J Diabetes research
26. clinical practice 157:107843.
27. Schroten-Loef C, Verhoeven R, de Hingh I, van de Wouw A, van Laarhoven H, Lemmens V (2018). Unknown primary carcinoma in the Netherlands: decrease in incidence and survival times remain poor between 2000 and 2012. Eur J Cancer 101:77-86.
28. Vajdic C, Perez-Concha O, Dobbins T, Ward RL, Schaffer AL, van Leeuwen M et al. (2019a). Demographic, social and lifestyle risk factors for cancer registry-notified cancer of unknown primary site (CUP). Cancer Epidemiol 60:156-161.
29. Vajdic C, Perez-Concha O, Rhee JJ, Dobbins T, Ward RL, Schaffer AL et al. (2019b). Health-related predictors of cancer registry-notified cancer of unknown primary site (CUP). Cancer Epidemiol 61:1-7.
30. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM (1990a). Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol 19(3):553-558.
31. Van den Brandt PA, Goldbohm RA, van 't Veer P, Volovics A, Hermus RJ, Sturmans F (1990b). A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol 43(3):285-295.
32. Van der Sanden G, Coebergh J-W, Schouten L, Visser O, Van Leeuwen F (1995). Cancer incidence in the Netherlands in 1989 and 1990: First results of the nationwide Netherlands Cancer Registry. Eur J Cancer 31(11):1822-1829.

CHAPTER 8

ADHERENCE TO THE WORLD CANCER RESEARCH FUND AND THE AMERICAN INSTITUTE FOR CANCER RESEARCH LIFESTYLE RECOMMENDATIONS FOR CANCER PREVENTION AND CANCER OF UNKNOWN PRIMARY RISK

Karlijn E.P.E. Hermans, Piet A. van den Brandt,
Caroline Loef, Rob L. H. Jansen, and Leo J. Schouten

Clin. Nutr.: 2022; 41: 526-535

Abstract

Background & Aims: The World Cancer Research Fund (WCRF) and American Institute for Cancer Research (AICR) updated their cancer prevention recommendations in 2018. Adherence to these recommendations has been associated with lower cancer risk and mortality. However, adherence in relation to Cancer of Unknown Primary (CUP) risk has not been studied. This study investigates whether adherence to the WCRF/AICR recommendations is associated with CUP risk.

Methods: Data from the prospective Netherlands Cohort Study on diet and cancer was used to measure adherence to the recommendations in relation to CUP risk. The cohort includes 120,852 participants (aged 55-69 years), who completed a self-administered questionnaire on cancer risk factors at baseline. Adherence was investigated with respect to body fatness, physical activity, plant foods, meat consumption and alcohol. Incident CUP cases were identified through record linkage to the Netherlands Cancer Registry and Dutch Pathology Registry. A follow-up of 20.3 years, resulted in 856 incident CUP cases and 3,911 subcohort members with complete information available for case-cohort analyses. Multivariable adjusted hazard ratios were estimated using proportional hazards models and were adjusted for age at baseline, sex, cigarette smoking (status, frequency, and duration) and total energy intake.

Results: Highest adherence appeared to be associated with decreased CUP risk in the age-sex adjusted model (HR: 0.76, 95% CI: 0.62-0.92). After additional adjustment for cigarette smoking (status, frequency, and duration), the association attenuated and was no longer statistically significant. No multiplicative interactions were observed between sex nor smoking status and overall adherence in relation to CUP.

Conclusion: Within this cohort, highest adherence to the WCRF/AICR recommendations is not statistically significantly associated with decreased CUP risk after multivariable adjustment.

Introduction

Cancer of Unknown Primary (CUP) is a metastasized malignancy with no identifiable primary tumor origin during life [1, 2]. The global CUP incidence has decreased over the last 10-20 years, and currently reaches 1-2% of all cancers [3]. In 2018, approximately 1,300 incident cases were registered by the Netherlands Cancer Registry (NCR) in the Netherlands [2]. CUP is a complex disease with a bleak prognosis due to the presence of metastases and the difficulty in identifying its primary tumor origin. The median survival for CUP patients ranges between three to ten months dependent on its histology [4].

CUP aetiology studies, including previous investigations in the Netherlands Cohort Study on diet and cancer (NLCS), reported modifiable risk factors such as cigarette smoking [5-8] and alcohol consumption [6-8] to be associated with increased CUP risk. For processed meat consumption, a moderate increased CUP risk was observed [8, 9]. In contrast, no association has been found between red meat consumption and CUP risk [8, 9], body mass index (BMI) [7, 8, 10], physical activity [8, 10], or vegetable and fruit consumption [8, 11] in relation to CUP risk.

The World Cancer Research Fund (WCRF) and the American Institute for Cancer Research (AICR) updated (2018) their cancer prevention recommendations with respect to modifiable lifestyle factors such as diet, nutrition, and physical activity [12]. It has been shown that adhering to these recommendations is associated with a lower risk of general and site-specific cancer, as well as reduced total and cancer-specific mortality [13-15].

Until now, only a few studies have investigated CUP aetiology, and to the best of our knowledge, no study examined whether adherence to the WCRF/AICR cancer prevention recommendations is associated with a decreased CUP risk. In general, the identification of modifiable lifestyle factors that are associated with a disease, may guide primary prevention in order to reduce its occurrence [16, 17]. Therefore, this study investigates whether adherence to the lifestyle recommendations regarding body fatness, physical activity, plant food consumption, meat consumption (red and processed meats), and alcohol is associated with CUP risk. In order to do so and to study the impact of the risk factors in detail, exposures were investigated as an overall adherence, as well as individual component adherence.

Materials & Methods

Study design and population

Adherence to the WCRF/AICR lifestyle recommendations in association with CUP risk was studied using data from the NLCS. According to the Population Intervention Comparison Outcome Study design, the population consist of participants who were followed-up for cancer incidence within the Netherlands Cohort Study on diet and cancer from September 1986 until December 2006. The population includes 120,852 participants aged 55-69 years at baseline in 1986, who originated from 204 Dutch municipal population registries [18]. The intervention group includes highest adherence to the WCRF/AICR lifestyle recommendations on cancer prevention with respect to the following exposures: body fatness, physical activity, plant foods, meat and alcohol consumption, whereas the control group measures lowest adherence to the abovementioned lifestyle components. The study design to measure the exposure-outcome relation is a prospective cohort, for which efficient data processing and analysis were achieved by applying a case-cohort approach. Subsequently, incident cancer cases were derived from the full cohort, while the number of person-years at risk for the full cohort was estimated from a subcohort of 5,000 participants who were randomly sampled from the full cohort at baseline [18].

Outcome measure

CUP is defined as a metastasized epithelial malignancy with no identifiable primary tumor origin after cytological and/or histological verification during a patient's lifetime. This definition only includes epithelial malignancies according to the International Classification of Diseases for Oncology version 3: M-8000 - M-8570. For this reason, cases with an unknown primary tumor origin and a histology of sarcoma, lymphoma, mesothelioma and melanoma were excluded from the analyses.

Follow-up

Participants were followed up for 20.3 years (from 17 September 1986 until 31 December 2006). Incident CUP cases were identified through annual record linkage with the Netherlands Cancer Registry (NCR) and the Dutch Pathology Registry (PALGA) [19]. Participants were removed from the analyses if there was 1) incomplete data on body fatness, physical activity, plant foods, meat consumption (red and processed meats), and alcohol, or if confounder data were missing; 2) inconsistent

dietary data concerning plant and meat consumption; 3) evidence that participants reported a history of cancer (except for skin cancer) at baseline. As a result, 856 CUP cases with a microscopical confirmation and epithelial histology remained, and a total of 3,911 subcohort members were available for analyses (see Figure 1).

Data collection

Participants of the NLCS completed a mailed, self-administered questionnaire on dietary habits and other cancer risk factors at baseline in 1986. Details on foods and beverages were evaluated for their validity and reproducibility [20, 21]. Dietary intake was recorded over three periods of three consecutive days each, to represent consumption patterns of vegetables, fruits, and meats during three seasons in the Netherlands. To evaluate the validity of the questionnaire, three parameters: 1) ratio of FFQ to record nutrient intake, 2) correlation coefficient, and 3) the distribution of mean nutrient intakes were compared, and were deemed to be accurate for measuring intake of food groups and nutrients [20]. No validation studies were conducted for measuring BMI, physical activity or smoking behavior. The following WCRF/AICR lifestyle recommendations were measured to study adherence in relation to CUP risk: 'be a healthy weight', 'be physically active', 'eat a diet rich in wholegrains, vegetables, fruit and beans', 'limit consumption of red and processed meat', and 'limit alcohol consumption'. In order to measure being a healthy weight, self-reported data on BMI at baseline and BMI at age 20 years was used in which weight at baseline and weight at age 20 years were divided by height at baseline squared (kg/m^2). Change in BMI since age 20 years, representing weight gain, was calculated as BMI at baseline minus BMI at age 20 years [10]. Non-occupational physical activity was calculated based on questions regarding gardening, cycling and walking, and sports/physical exercise in minutes per day [10]. Plant-based foods were measured using data on dietary fiber intake in grams per day, and the amounts of vegetables and fruits consumed in grams per day [11]. The Spearman correlation coefficients for measuring the validity of total vegetable consumption and total fruit consumption were 0.38 and 0.60, respectively, compared to results of a 9-day diet record [20]. Questions regarding meat consumption specifically addressed red meat (unprocessed) and processed meat consumption [9]. Red meat includes: beef, pork, minced meat (beef and pork), liver, and other meat (e.g., horsemeat, lamb). Processed meat includes: ham, bacon, smoked beef or pork loin roll, and other sliced cold meats (e.g., sausages). The Spearman correlation

coefficients for the validity of red meat and processed meat as investigated by the questionnaire were 0.46 and 0.54, respectively, compared to the results of the 9-day diet record [20]. Data on alcohol consumption was obtained through questions on the consumption of beer, red wine, white wine, sherry, other fortified wines, liqueurs and liquor [6]. Mean daily alcohol consumption was calculated by using the computerized Dutch food composition table [22]. Based on pilot study data, standard glass sizes were defined as 200 ml for beer, 105 ml for wine, 80 ml for sherry, and 45 ml for both liqueurs and liquor, corresponding to 8, 10, 11, 7, and 13 grams of ethanol, respectively. The Spearman correlation coefficient between alcohol consumption as assessed by the questionnaire and that estimated by the 9-day diet record was 0.89 for all subjects and 0.85 for alcohol consumers [20]. The questionnaire also included questions on baseline smoking status (cigarette, cigar, or pipe), smoking frequency, and the ages at first exposure and last (if stopped) exposure to smoking [6]. Participants who indicated that they had never smoked were considered never smokers. To avoid collinearity problems, smoking frequency and duration were centered as proposed by Leffondré et al. [23].

WCRF/AICR sumscore

Adherence to the WCRF/AICR cancer prevention recommendations was measured using methods similar to those applied in previous adherence studies [13, 14]. Five of the ten recommendations (be a healthy weight; be physically active; eat a diet rich in wholegrains, vegetables, fruits and beans; limit consumption of red and processed meat; and limit alcohol consumption) were used to calculate an overall adherence score (**see Table 1**). The remaining five recommendations (limit consumption of 'fast foods' and other processed foods high in fat, starches or sugars; limit consumption of sugar sweetened drinks; do not use supplements for cancer prevention; for mothers: breastfeed your baby if you can; after a cancer diagnosis: follow our recommendations if you can) were not included in the analysis either because they were not (optimally) measured at baseline in this cohort, or do not apply to the research question.

For each of the five recommendations with available data, a maximum of 1 point could be obtained, therefore, the overall adherence score ranged from 0 to 5 points. Overall adherence was measured using three cut-off categories which were distributed as evenly as possible on the basis of the subcohort: lowest adherence: 0-<2 points (reference group), medium adherence: 2-<3.5 points and

highest adherence: ≥3.5 points. For the recommendations be a healthy weight, be physically active, limit consumption of red and processed meat, and limit alcohol consumption; participants could receive a half point for partially complying with the recommendation, or a maximum of 1 point for fully complying with the recommendation. The recommendation 'eat a diet rich in wholegrains, vegetables, fruits and beans' included two sub-recommendations which involved: 1) fiber intake and 2) vegetable and fruit consumption; for each sub-recommendation, 0.25 points could be received for partially complying with the sub-recommendation, or a maximum of 0.5 points for fully complying with the sub-recommendation. Using cut-off values as proposed by the WCRF/AICR, most NLCS-participants clustered in the highest adherence levels for physical activity and plant food consumption, whereas the majority of participants clustered in the lowest adherence category for meat consumption. Therefore, to measure variation in adherence, we used 30-minute increment categories for physical activity; and tertiles as cut-off values for plant food and meat consumption, representing lowest, middle and highest adherence.

Ethics

The institutional review boards of the Netherlands Organization for Applied Scientific Research TNO (Zeist) and Maastricht University (Maastricht) approved the execution of the NLCS. Participants agreed to be included into the cohort and follow-up by returning the questionnaire they completed.

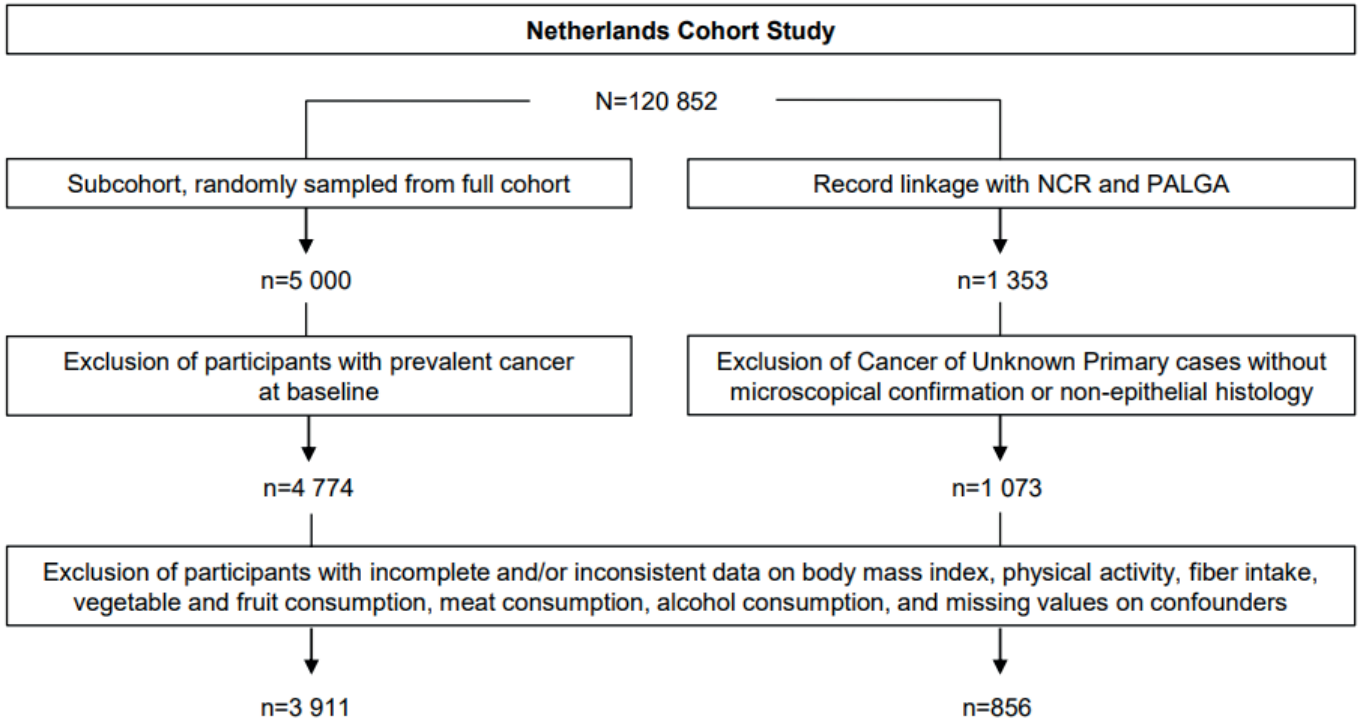


Figure 1 Flow diagram of subcohort members and Cancer of Unknown Primary cases in the Netherlands Cohort Study on whom analyses are based

Table 1 Operationalization adherence scores based on the WCRF/AICR lifestyle recommendations (2018) in the Netherlands Cohort Study

			Points
1.	**Be a healthy weight.** Keep your weight within the healthy range and avoid weight gain in adult life		
*	Ensure that body weight during childhood and adolescence projects towards lower end of the healthy adult BMI range	Insufficient data available	
*	Keep your weight as low as you can within the healthy range throughout life	BMI (kg/m^2):	
		18.5 – 24.9	1
		25.0 – 29.9	0.5
		<18.5 or >30	0
*	Avoid weight gain (measured as body weight or waist circumference) throughout adulthood	*Change in BMI (kg/m^2): **	
		0-<4	*0.5*
**	Weight gain will only be studied in a sensitivity analysis	*4-<8*	*0.25*
		≥8	*0*
2.	**Be physically active.** Be physically active as part of everyday life – walk more and sit less		
*	Be at least moderately physically active and follow or exceed national guidelines	PA (min/day)	
		≥60 min/d	1
		30-<60 min/d	0.5
		<30 min/d	0
*	Limit sedentary habits	Insufficient data available	
3.	**Eat a diet rich in wholegrains, vegetables, fruits and beans.** Make wholegrains, vegetables, fruit, and pulses such as beans and lentils a major part of your usual daily diet		
*	Consume a diet that provides at least 30 grams per day of fibre	Tertiles fibre intake	
		Highest	0.5
		Middle	0.25
		Lowest	0
*	Eat diet high in all types of plant foods including at least five portions (400 grams in total) of a variety of non-starchy vegetables and fruit every day	Tertiles vegetable and fruit consumption	
		Highest	0.5
		Middle	0.25
		Lowest	0

			Points
*	Include in most meals foods containing wholegrains, non-starchy vegetables, fruit and pulses (legumes) such as beans and lentils	Insufficient data available	
*	If you eat starch roots and tubers as staple foods, eat non-starchy vegetables, fruits and pulses regularly too if possible	Insufficient data available	
4.	Limit consumption of fast foods and other processed foods high in fat, starches or sugars. **Limiting these foods helps control calorie intake and maintain a healthy weight**		
*	Limit consumption of fast foods and other processed foods high in fat, starches or sugars – including fast foods; many pre-pared dishes, snacks, bakery foods and desserts; and confectionery (candy)	Insufficient data available	
5.	**Limit consumption of red and processed meat.** Eat no more than moderate amounts of red meat, such as beef, pork and lamb. Eat little, if any, processed meat		
*	Limit consumption to no more than 3 portions per week of red meat (350-500 grams cooked weight of the meat). Consume very little processed meat	Tertiles meat consumption	
		Highest	1
		Middle	0.5
		Lowest	0
6.	**Limit consumption of sugar sweetened drinks.** Drink mostly water and unsweetened drinks	Insufficient data available	
7.	**Limit alcohol consumption.** For cancer prevention, it's best not to drink alcohol	Ethanol intake (g/d)	
		0	1
		>0 to ≤10	0.5
		>10	0
8.	**Do not use supplements for cancer prevention**. Aim to meet nutritional needs through diet alone	Not measured in relation to cancer prevention	
9.	**For mothers: breastfeed your baby if you can.** Breastfeeding is good for both mother and baby	Not measured in this population	
10.	**After a cancer diagnosis: follow our recommendations, if you can.** Check with your health professional what is right for you	Not applicable to the research question	

Statistics

Person-years at risk were calculated from baseline (17 September 1986) until CUP diagnosis, death, emigration, loss to follow-up, or end of follow-up (31 December 2006), whichever occurred first. General characteristics for subcohort members and CUP cases were presented as frequencies (percentages) for categorical variables and means including standard deviations for continuous variables.

Predefined confounders included age at baseline (years; continuous); and sex (male/female), cigarette smoking status (never, ex, current); cigarette smoking frequency (centered; number of cigarettes smoked per day); cigarette smoking duration (centered; number of years smoking), and total energy intake (kcal/day; continuous). Potential confounders included socio-economic status (highest level of education); doctor's diagnosis of diabetes (yes/no); and history of cancer in a first degree relative (yes/no). Variables were considered a confounder if they changed the HR by >10%. Accordingly, none of the potential confounders were added in the final model.

Age- and sex-adjusted, and multivariable-adjusted hazard ratios (HRs) with 95% confidence intervals (CIs) were estimated using the Cox proportional hazards models. Standard errors were calculated using the robust Huber-White sandwich estimator to account for additional variance introduced by sampling from the full cohort [24]. The proportional hazards assumption was tested using the scaled Schoenfeld residuals [25]. In case the assumption was violated, log-minus-log (LML) survival curves were inspected visually. If necessary, a time-varying covariate (TVC) for that variable was added to the model. Ordinal exposure variables were fitted as continuous variables in trend analyses. Wald tests and cross-product terms were applied to evaluate potential multiplicative interaction between sex or smoking status in relation to overall adherence and CUP risk. All analyses were conducted using Stata software (version 15). *P* values were considered statistically significant if $p < 0.05$.

An additional analysis was conducted in which we tested whether the exclusion of alternating individual components affected the overall adherence outcome. For this purpose, Z-scores were calculated per one-point increment to enable standardized comparisons between the models. In another analysis, weight gain was included to the model as part of the recommendation on body fatness. We did not include weight gain in the overall model, as the inclusion of this variable results in fewer CUP cases (n=714) and fewer subcohort members (n=3,345) because of additional missing values. For this analysis, participants with complete data on BMI at baseline and weight gain could respectively receive a maximum of 0.5 points for complying

with the recommendation. In a sensitivity analysis, we restricted the analysis only to histologically verified CUP cases as those cases were more likely to meet stringent CUP definitions, and/or were more likely to have undergone extensive diagnostic investigation(s) than cytologically verified cases; due to other factors that may have influenced the decision to refrain from further diagnostic investigation such as age, comorbidities, performance status, localization of metastasis, or the patient's decision. Lastly, another sensitivity analysis was conducted which saw the first two-years of follow-up excluded from the analysis so as to check for potential reverse causality bias as a result of preclinical cancer at baseline.

Results

Statistical analyses are based on 856 incident CUP cases and 3,911 subcohort members with complete data on body fatness, physical activity, plant foods, meat consumption, alcohol, and confounder variables.

On average CUP cases were aged 62 years at baseline, while subcohort members were aged 61 years at baseline (**see Table 2**). CUP cases were predominantly men (63.2%), whereas the sex-distribution was more equal in the subcohort members (50.4% men). The majority of CUP cases had a BMI between 18.5 kg/m^2 and <25 kg/m^2 at baseline (56.7%), which was slightly more than in the subcohort members (54.0%). Most CUP cases were physically active >60 minutes per day (49.0%), which was comparable in the subcohort members (49.1%). We observed that male CUP cases consumed a slightly lower amount of fiber compared to male subcohort members (27.7 g/day versus 28.5 g/day); that observation held for female CUP cases in comparison to female subcohort members (24.9 g/day versus 25.2 g/day). Male CUP cases also consumed slightly lower amounts of vegetables and fruits compared to male subcohort members (330.2 g/day versus 341.6 g/day), whereas female CUP cases and female subcohort members consumed equal amounts (386.1 g/day versus 386.6 g/day). Both male and female CUP cases ate more red and processed meats in comparison to male and female subcohort members (111.6 g/day and 95.9 g/day versus 109.7 g/day and 91.3 g/day, respectively. Most CUP cases consumed >10 g of ethanol per day (44.6%), whereas most subcohort members consumed 0-≤10 g of ethanol per day (41.5%).

The mean WCRF/AICR adherence score was 2.71 (sd: 0.87), ranging from 0.50 to 4.75 in the CUP cases, whereas it was 2.82 (sd: 0.89), ranging from 0.75 to 4.75 points in the subcohort. Using the Cox proportional hazards models, no multiplicative interaction was found between sex and overall adherence in relation to CUP risk ($P_{interaction}$ = 0.25),

nor between smoking status and overall adherence in relation to CUP risk ($P_{interaction}$ = 0.74). Consequently, no sex- or smoking status-stratified results will be presented. Adherence to the WCRF/AICR lifestyle recommendations was measured as an overall adherence as well as individual component adherence (**see Table 3 and 4**). In the overall adherence model, highest adherence to the lifestyle recommendations appeared to be associated with a statistically significant decreased CUP risk compared to lowest adherence in the age-sex adjusted analysis (age-sex adjusted HR: 0.76, 95% CI: 0.62-0.92). However, after additionally adjusting for cigarette smoking (status, frequency, and duration) and total energy intake, the association attenuated and was no longer statistically significant (multivariable adjusted HR: 0.87, 95% CI: 0.70-1.08). Adherence with respect to individual components concerning body fatness, physical activity, and intake of plants foods did not appear to be associated with CUP risk. CUP cases with the highest adherence to the recommendations for meat (red and processed meats) and alcohol consumption appeared to have the lowest CUP risk (multivariable adjusted HR: 0.80, 95% CI: 0.65-0.98, P_{trend} = 0.03 & multivariable adjusted HR: 0.74, 95% CI: 0.59-0.93, P_{trend} = 0.01, respectively).

In an additional analysis, we compared the overall adherence outcome per model after excluding alternating individual components. The HRs of the models were compared per one-point increment based on Z-scores. We observed that after comparing these HRs, the decreased CUP risk became stronger after excluding BMI from the overall sumscore (multivariable adjusted HR: 0.89, 95% CI: 0.83-0.97), whereas the association did not change after excluding physical activity, but attenuated after excluding plant foods-, meat-, and alcohol consumption, respectively (see Figure 2). In another analysis, weight gain was included to the model as part of the recommendation on body fatness. Although this enabled us to study body fatness in greater detail, the inclusion resulted in fewer CUP cases (n=714) and fewer subcohort members (n=3,345) available. By applying this smaller number of CUP cases and subcohort members, we observed that CUP risk was again decreased in the highest adherence category (multivariable adjusted HR: 0.73, 95% CI: 0.53-1.01) compared to the lowest adherence category (data not shown). After including the weight gain variable into the model, we again detected similar findings to those in the smaller case mix and subcohort. The results of the sensitivity analysis when restricted to histologically verified CUP cases alone do not differ substantially from the results of the overall analysis (data not shown). Findings of another sensitivity analysis, in which the first two years of follow-up were excluded, also revealed comparable results to those of the overall analysis (data not shown).

Table 2 General characteristics of Cancer of Unknown Primary cases and subcohort members in the Netherlands Cohort Study

	Subcohort members (n=3911)		Cancer of Unknown Primary cases (n=856)	
Characteristic	**n**	**(%)**	**n**	**(%)**
Age at baseline (years), mean (SD)				
Overall	61.3 (4.2)		62.0 (4.1)	
Sex				
Men	1971	50.4	541	63.2
Women	1940	49.6	315	36.8
BMI (kg/m²) at baseline				
<18.5	37	1.0	1	0.1
18.5-<25	2110	54.0	485	56.7
25-<30	1518	38.8	317	37.0
≥30	246	6.3	53	6.2
Non-occupational physical activity (min/day)				
<30	749	19.2	165	19.3
30-60	1241	31.7	272	31.8
>60	1921	49.1	419	49.0
Total fiber intake (g/day), mean (SD)				
Men	28.5 (7.2)		27.7 (6.7)	
Women	25.2 (7.4)		24.9 (5.9)	
Total vegetable and fruit consumption (g/day), mean (SD)				
Men	341.6 (149.1)		330.2 (144.0)	
Women	386.6 (151.5)		386.1 (145.2)	
Total meat consumption (g/day), mean (SD)				
Men	109.7 (46.5)		111.6 (43.9)	
Women	91.3 (41.7)		95.9 (40.3)	

	Subcohort members (n=3911)		Cancer of Unknown Primary cases (n=856)	
Characteristic	**n**	**(%)**	**n**	**(%)**
Ethanol intake (g/day) [1]				
Abstainers	909	23.2	151	17.6
0-≤10	1622	41.5	323	37.7
>10	1380	35.3	382	44.6
Cigarette smoking status				
Never smokers	1362	34.8	226	26.4
Ever smokers	2549	65.2	630	73.6
Level of education (years of education)				
Primary	1074	27.6	222	26.2
Lower vocational	839	21.6	169	19.9
Secondary and medium vocational	1408	36.2	323	38.0
University and higher vocational	572	14.7	135	15.9
Diabetes				
Yes	130	3.3	28	3.3
First grade family history of cancer				
Yes	1787	45.7	416	48.6

[1] In consumers only

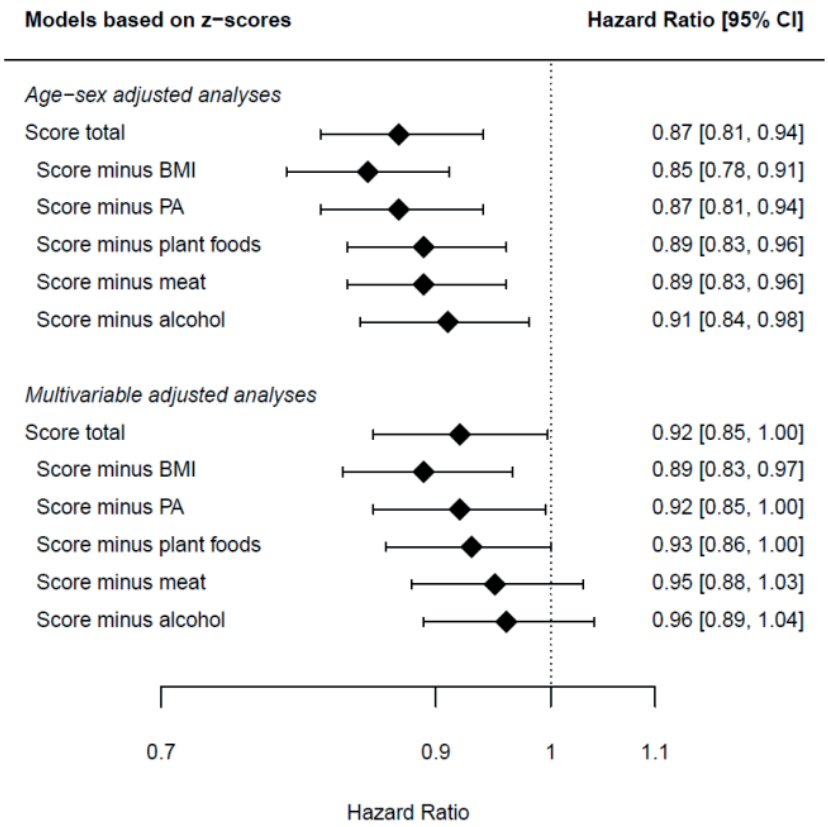


Figure 2 Overall WCRF/AICR adherence scores (per one-point increment based on Z-scores) applied in the Netherlands Cohort Study – individual component exclusion

Notes

-Scoring system is based on the 2018 WCRF/AICR lifestyle recommendations on cancer prevention, scores are based on the National Cancer Institute operationalization and the distribution of subcohort members in the NLCS-cohort.

-Age and sex-adjusted analyses were adjusted for age at baseline (years), and sex.

-Multivariable adjusted analyses were adjusted for age at baseline (years), sex, cigarette smoking status (never/ever; centered), cigarette smoking frequency (continuous; centered), cigarette smoking duration (continuous; centered), and total energy intake (kcal/day).

-BMI = body mass index, PA = physical activity, plant foods = fiber intake & vegetable and fruit consumption, meat = red and processed meat consumption, alcohol = alcohol consumption.

Table 3 Overall Adherence to the World Cancer Research Fund (WCRF) and American Institute for Cancer Research (AICR) lifestyle recommendations on cancer prevention and adherence in relation to Cancer of Unknown Primary (CUP) risk in the Netherlands Cohort Study [1]

		Subcohort members			Cancer of Unknown Primary cases							
		Adherence		Person time at	Adherence		Age and sex-adjusted [2]			Multivariable adjusted [3]		
		n	%	risk (years)	*n*	%	HR	95% CI		HR	95% CI	
Overall WCRF/AICR score												
<2.0	lowest adherence	1196	30.6	20 082	282	32.9	1	Reference		1	Reference	
2.0-<3.5	medium adherence	1583	40.5	26 459	368	43.0	1.03	(0.86-	1.22)	1.08	(0.90-	1.31)
≥3.5	highest adherence	1132	28.9	19 508	206	24.1	0.76	(0.62-	0.92)	0.87	(0.70-	1.08)
	p for trend [4]						0.93			0.99		

[1] Scoring system is based on the 2018 WCRF/AICR lifestyle recommendations on cancer prevention, scores are based on the National Cancer Institute operationalisation and the distribution of subcohort members in the NLCS-cohort.

[2] Analyses were adjusted for age at baseline (years), and sex.

[3] Analyses were adjusted for age at baseline (years), sex, cigarette smoking status (never/ever; centered), frequency (continuous; centered), duration (continuous; centered), and total energy intake (kcal/day).

[4] Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

Table 4 Components of the World Cancer Research Fund (WCRF) and American Institute for Cancer Research (AICR) lifestyle recommendations on cancer prevention and adherence in relation to Cancer of Unknown Primary (CUP) risk in the Netherlands Cohort Study [1]

		Score		Subcohort members (n=3911)			Cancer of Unknown Primary cases (n=856)					
				Adherence		Person time at	Adherence		Age and sex-adjusted [2]		Multivariable adjusted [3]	
				n	%	risk (years)	*n*	%	HR	95% CI	HR	95% CI
1	**Be a healthy weight**		BMI (kg/m²)									
		1	18.5-24.9	2110	54.0	35 882	485	56.7	1	Reference	1	Reference
		0.5	25.0-29.9	1518	38.8	25 459	317	37.0	0.88	(0.75- 1.03)	0.86	(0.72- 1.01)
		0	<18.5 or >30	283	7.2	4 707	54	6.3	0.99	(0.73- 1.37)	0.92	(0.66- 1.27)
			p for trend [4]						0.14		0.16	
2	**Be physically active**		PA (min/day)									
		1	≥60	1921	49.1	32 557	419	49.0	1	Reference	1	Reference
		0.5	30-<60	1241	31.7	21 394	272	31.8	1.00	(0.84- 1.19)	0.98	(0.82- 1.17)
		0	<30	749	19.2	12 098	165	19.3	1.14	(0.93- 1.40)	1.07	(0.86- 1.32)
			p for trend [4]						0.27		0.66	
3	**Eat a diet rich in wholegrains, vegetables, fruits, and beans**		Tertiles fiber intake									
		0.5	Highest intake level	1304	33.3	22 064	260	30.4	1	Reference	1	Reference
		0.25	Middle intake level	1303	33.3	22 436	294	34.4	1.11	(0.92- 1.34)	1.04	(0.86- 1.27)
		0	Lowest intake level	1304	33.3	21 548	302	35.3	1.22	(1.02- 1.47)	1.05	(0.86- 1.28)
			p for trend [4]						0.03		0.64	
			Tertiles vegetable and fruit consumption									
		0.5	Highest intake level	1303	33.3	22 284	262	30.6	1	Reference	1	Reference
		0.25	Middle intake level	1303	33.3	22 094	292	34.1	1.15	(0.95- 1.38)	1.10	(0.91- 1.34)
		0	Lowest intake level	1305	33.4	21 671	302	35.3	1.23	(1.02- 1.49)	1.09	(0.90- 1.33)
			p for trend [4]						0.03		0.39	

		Score		Subcohort members (n=3911)			Cancer of Unknown Primary cases (n=856)						
				Adherence		Person time at	Adherence		Age and sex-adjusted [2]			Multivariable adjusted [3]		
				n	%	risk (years)	*n*	%	HR	95% CI		HR	95% CI	
4	**Limit consumption of red and processed meat**		Tertiles meat consumption											
		1	Lowest intake level	1306	33.4	22 057	254	29.7	0.78	(0.65-	0.94)	0.80	(0.65-	0.98)
		0.5	Middle intake level	1303	33.3	21 874	292	34.1	0.93	(0.77-	1.11)	0.94	(0.77-	1.15)
		0	Highest intake level [5]	1302	33.3	22 117	310	36.2	1	Reference		1	Reference	
			p for trend [4]						0.01			0.03		
5	**Limit alcohol consumption**		Ethanol intake (g/day)											
		1	0	909	23.2	15 145	151	17.6	0.72	(0.58-	0.89)	0.74	(0.59-	0.93)
		0.5	>0-≤10	1622	41.5	27 979	323	37.7	0.80	(0.67-	0.94)	0.88	(0.74-	1.06)
		0	>10 [5]	1380	35.3	22 925	382	44.6	1	Reference		1	Reference	
			p for trend [4]						0.001			0.01		

1 Scoring system is based on the 2018 WCRF/AICR lifestyle recommendations on cancer prevention, scores are based on the National Cancer Institute operationalisation and the distribution of subcohort members in the NLCS-cohort.

2 Analyses were adjusted for age at baseline (years), and sex.

3 Analyses were adjusted for age at baseline (years), sex, cigarette smoking status (never/ever; centered), frequency (continuous; centered), duration (continuous; centered), and total energy intake (kcal/day).

4 Tests for dose-response trends were assessed by fitting ordinal variables as continuous terms in the Cox proportional hazards model.

5 Highest intake levels of meat and alcohol consumption represent the lowest adherence to the WCRF/AICR lifestyle recommendations.

Discussion

This large prospective cohort study, to our knowledge, is the first study to have investigated adherence to the WCRF/AICR lifestyle recommendations regarding body fatness, physical activity, plant foods, meat consumption, and alcohol in relation to CUP risk. The overall adherence model indicates that adherence was no longer statistically significant after additional adjustment for cigarette smoking (status, frequency, and duration) and total energy intake. No multiplicative interactions were observed between sex nor smoking status and overall adherence in relation to CUP. Meat (red and processed meats) and alcohol consumption appear to be the drivers for the overall adherence effect, as highest adherence for these exposures was significantly associated with decreased CUP risk. Adherence to the recommendations with respect to body fatness, physical activity or intake of plant foods was not associated with CUP risk.

To study the general clustering of health-related behavior, overall adherence was investigated in which highest adherence to the lifestyle recommendations appeared to be associated with a significantly decreased CUP risk compared to the lowest adherence category in the age-sex adjusted analysis (HR: 0.76, 95% CI: 0.62-0.92). Yet, the association attenuated and was no longer statistically significant after additionally adjusting for cigarette smoking (status, frequency, and duration) and total energy intake. To check whether the attenuation derived from the influence by smoking behavior and/or total energy intake, we compared estimates after individually correcting for these variables. After correcting for total energy intake alone, the decreased CUP risk persisted and remained statistically significant. After correcting for smoking behavior alone, the association attenuated and was no longer statistically significant. Thus, smoking appears to influence the overall adherence association within this cohort. In addition, we have seen that men and women who were never smokers had higher mean WCRF/AICR adherence scores compared to ex- and current smokers which implies that never smokers generally have a healthier lifestyle in the NLCS.

The European Prospective Investigation into Cancer and Nutrition (EPIC) study assessed adherence to the WCRF/AICR lifestyle recommendations in relation to total and subsequent cancer risk. Its authors concluded that individuals in European populations who complied were less likely to develop various types of cancers than individuals who did not comply [13]. For total cancer, they reported that adherence was associated with

a statistically significant reduced risk (multivariable adjusted HR: 0.95, 95% CI: 0.93-0.97) which was also adjusted for smoking frequency and duration, as well as total energy intake. Other studies have also demonstrated that adopting a healthy pattern of diet (including foods and beverages with a relatively high concentration of vitamins and minerals), nutritional health (without excessive fats, added sugars or refined starches) and physical activity is beneficial for increasing longevity as well as protecting against both cancers overall and other noncommunicable diseases [13, 14, 26].

With respect to the individual components, epidemiological studies have investigated CUP etiology with respect to body fatness, physical activity, plant foods, meat consumption and alcohol. These studies include the EPIC cohort (651 CUP cases) [7], an Australian prospective cohort (327 CUP cases compared to two sets of controls) [8], a Swedish case-control study (447 CUP cases) [5], and our NLCS (867-963 CUP cases dependent on the availability of exposure data and number of missing values) [6, 9-11]. In the current study, we observed that highest adherence to the recommendations with respect to meat (red and processed meats) and alcohol consumption is significantly associated with decreased CUP risk, which is in agreement with the studies that found meat and alcohol consumption to increase CUP risk [7, 8]. Inversely, we noted that adherence was not associated with body fatness, physical activity or intake of plant foods and CUP risk. These results are also in line with the findings of the previous studies that found no association between BMI [7, 8], physical activity [8], or vegetable and fruit consumption [8] in relation to CUP. After excluding alternating individual components to compare the overall adherence outcome per model (HRs of the models per one-point increment based on Z-scores), we found that the decreased CUP risk became stronger after excluding BMI, while the risk remained the same after excluding physical activity, but the association attenuated for plant foods-, meat- and alcohol consumption. These abovementioned findings suggest that meat and alcohol consumption play an important role in the overall adherence effect within our cohort as highest adherence to the recommendations for both components is significantly associated with decreased CUP risk. It is unlikely that BMI, physical activity, or plant foods consumption affected the overall adherence outcome, as no substantial effects were observed from these components.

The strengths of the NLCS are its prospective cohort design that includes an extensive cohort of 120,852 participants who were followed-up for 20.3 years. In this study, we were able to investigate 856 incident CUP cases, which is a higher number than other studies were able to investigate with respect to CUP etiology. The comprehensive

questionnaire of the NLCS allowed us to study the most important lifestyle factors for reducing cancer risk as on overall adherence, as well as individual component adherence. In addition, there was a wide availability of confounder data, which showed to be of great importance in our multivariable adjusted analysis, as smoking behavior appeared to attenuate the statistically significant decreased CUP risk that was observed in the age-sex adjusted analysis. Completeness of record linkage with the NCR and PALGA was at least 96% for cancer follow-up, and vital status follow-up was complete for almost 100% after 20.3 years, thereby minimizing selection bias [27]. Information on incident CUP cases was obtained from the NCR, which included specific details from both pathology and clinical reports [28]. Registry clerks from the NCR applied uniform coding rules when entering data based on medical files. Therefore, we were able to analyze incident CUP cases with a consistent disease definition. Certain limitations need to be acknowledged. For example, five of the WCRF/AICR recommendations could not be included in our sumscore. Two of the recommendations are not included as: 1) breastfeeding, was not measured within the NLCS, and 2) after a cancer diagnosis, does not apply to the research question. The remaining three recommendations: 3) limiting the consumption of fast foods and other processed foods high in fat, starches or sugar, 4) limit the consumption of sugar sweetened drinks, were not adequately measured within the NLCS, these dietary habits were not common in the cohort, and 5) do not use supplements for cancer prevention, only included a low percentage of users, and reasoning for utilisation was unclear. The updated supplement recommendation states to not use supplements for cancer prevention, unfortunately we do not have data available to check as to why the supplements were used and therefore whether the reasoning was cancer prevention or lifestyle in general. Notwithstanding, the five recommendations with available data were extensively measured in our cohort which made it possible to study various lifestyle components at both an overall adherence level as well as an individual component adherence level, something which previous epidemiological studies had not conducted for CUP. The WCRF/AICR's 2018 Continuous Update Project has summarized that the most important factors for reducing cancer risk are to avoid smoking, to maintain a healthy weight throughout life by consuming a healthy diet, and being physically active. Since we were able to study these most important factors, we believe that our estimation is an adequate reflection of measuring adherence to the lifestyle recommendations in relation to CUP risk. In addition, we observed that for physical activity and plant food consumption, most participants within our cohort complied with highest adherence levels, whereas the majority of participants consumed more meats than recommended. Consequently,

variation in the cohort would have been reduced. Therefore, to measure the variation in adherence optimally, we used 30-minute increment categories for physical activity; and tertiles as cut-off values for plant and food consumption, representing lowest, middle and highest adherence to a healthy behavior. Furthermore, unfortunately, no validation studies were conducted for measuring BMI, physical activity or smoking behavior. For anthropometric measures, recall bias may have occurred as weight and height were asked at baseline in 1986. However, other studies have concluded that self-reported recall of anthropometric measures in early life is highly correlated with prospectively collected data [29]. With respect to physical activity, this study examined self-reported non-occupational physical activity as an indicator for exercise behavior. This self-reported measurement may have attenuated the association [30]. Smoking behavior was measured through various components such as smoking status (cigarette, cigar, or pipe), smoking frequency, and the ages at first exposure and last (if stopped) exposure to smoking. A review on the validity of self-reported smoking has concluded that it is not expected that there are major differences in self-reported smoking behavior due to the form of biochemical validation [31].

Researchers from the National Cancer Institute, WCRF, AICR, and the WCRF/AICR Continuous Update Project Expert Panel, together with other international researchers have agreed on a standardized scoring system in which each recommendation is weighted equally. However, there is no agreement on whether weighting should be equal within components, as this could result in underestimating their joint effect [15]. In addition, there are more dietary components than non-dietary components in the WCRF/AICR recommendations, which naturally gives greater weight to those dietary components.

Conclusion

In conclusion, highest adherence to the WCRF/AICR lifestyle recommendations was statistically significantly associated with decreased CUP risk in the age-sex adjusted analysis, while the association attenuated and was no longer significant after additionally adjusting for smoking behavior and total energy intake. Our additional analysis revealed that the attenuation derived from the correction for smoking behavior alone. Participants with highest adherence to the recommendations for meat (red and processed meats) and alcohol consumption were found to have statistically significantly decreased CUP risk. Adherence to the recommendations with respect to body fatness, physical activity or intake of plant foods was not associated with CUP risk.

References

1. Loffler H, Kramer A. Introduction. In: Loffler H, Kramer A, editors. Cancer of Unknown Primary Springer International Publishing Switzerland; 2016.
2. Schroten-Loef C, Verhoeven R, de Hingh I, van de Wouw A, van Laarhoven H, Lemmens V. Unknown primary carcinoma in the Netherlands: decrease in incidence and survival times remain poor between 2000 and 2012. Eur J Cancer. 2018;101:77-86.
3. Rassy E, Pavlidis N. The currently declining incidence of cancer of unknown primary. J Cancer Epidemiol. 2019;61:139-41.
4. Hemminki K. Epidemiology, Risk Factors, and Survival in CUP: Pointers to Disease Mechanisms. Cancer of Unknown Primary: Springer; 2016. p. 5-25.
5. Hemminki K, Chen B, Melander O, Manjer J, Hallmans G, Hemminki A. Smoking and body mass index as risk factors for subtypes of cancer of unknown primary. Int J Cancer. 2015;136:246-7.
6. Hermans KE, van den Brandt PA, Loef C, Jansen RL, Schouten LJ. Alcohol consumption, cigarette smoking and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study. Int J Cancer. 2021.
7. Kaaks R, Sookthai D, Hemminki K, Kramer A, Boeing H, Wirfalt E, et al. Risk factors for cancers of unknown primary site: Results from the prospective EPIC cohort. Int J Cancer. 2014;135:2475-81.
8. Vajdic C, Perez-Concha O, Dobbins T, Ward RL, Schaffer AL, van Leeuwen M, et al. Demographic, social and lifestyle risk factors for cancer registry-notified cancer of unknown primary site (CUP). Cancer Epidemiol. 2019;60:156-61.
9. Hermans KEPE, Van den Brandt PA, Loef C, Jansen RLH, LJ. S. Meat consumption and cancer of unknown primary (CUP) risk: results from The Netherlands cohort study on diet and cancer. Eur J Nutr. 2021.
10. Hermans KE, van den Brandt PA, Loef C, Jansen RL, Schouten LJ. Anthropometry, physical activity and cancer of unknown primary (CUP) risk: Results from the Netherlands cohort study. Cancer Epidemiol. 2020;69:101836.
11. Hermans KEPE, Van den Brandt PA, Loef C, Jansen RLH, LJ. S. Vegetable and fruit consumption and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study on Diet and Cancer. Submitted, personal communication.
12. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. Continuous Update Project Expert Report 20182018.
13. Romaguera D, Vergnaud AC, Peeters PH, van Gils CH, Chan DS, Ferrari P, et al. Is concordance with World Cancer Research Fund/American Institute for Cancer Research guidelines for cancer prevention related to subsequent risk of cancer? Results from the EPIC study. Am J Clin Nutr. 2012;96:150-63.
14. Vergnaud AC, Romaguera D, Peeters PH, van Gils CH, Chan DS, Romieu I, et al. Adherence to the World Cancer Research Fund/American Institute for Cancer Research guidelines and risk of death in Europe: results from the European Prospective Investigation into Nutrition and Cancer cohort study1,4. Am J Clin Nutr. 2013;97:1107-20.

15. Shams-White MM, Brockton NT, Mitrou P, Romaguera D, Brown S, Bender A, et al. Operationalizing the 2018 World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) cancer prevention recommendations: a standardized scoring system. J Nutrients. 2019;11:1572.
16. Wild CP, Bucher JR, De Jong BW, Dillner J, Von Gertten C, Groopman JD, et al. Translational cancer research: balancing prevention and treatment to combat cancer globally. J Natl Cancer Inst. 2015;107:1-5.
17. Dovjak M, Kukec A. Creating Healthy and Sustainable Buildings: An Assessment of Health Risk Factors: Springer Nature; 2019.
18. Van den Brandt PA, Goldbohm RA, van 't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol. 1990;43:285-95.
19. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol. 1990;19:553-8.
20. Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, et al. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin Nutr. 1994;48:253-65.
21. Goldbohm RA, van 't Veer P, van den Brandt PA, van 't Hof MA, Brants HA, Sturmans F, et al. Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. Eur J Clin Nutr. 1995;49:420-9.
22. NEVO-table. Dutch food composition table 1986–1987. Voorlichtingsbureau voor de voeding, The Hague, The Netherlands. 1986.
23. Leffondre K, Abrahamowicz M, Siemiatycki J, Rachet B. Modeling smoking history: a comparison of different approaches. American journal of epidemiology. 2002;156:813-23.
24. Barlow WE. Robust variance estimation for the case-cohort design. Biometrics. 1994:1064-72.
25. Lin DY, Wei L-J. The robust inference for the Cox proportional hazards model. J Am Stat Assoc. 1989;84:1074-8.
26. World Cancer Research Fund/American Institute for Cancer Research. Recommendations and public health and policy implications. Continuous Update Project Expert Report2018.
27. Goldbohm RA, Van den Brandt PA, Dorant E. Estimation of the coverage of Dutch municipalities by cancer registries and PALGA based on hospital discharge data. Tijdschr Soc Gezondheidsz 1994;72:80-4.
28. Van der Sanden G, Coebergh J-W, Schouten L, Visser O, Van Leeuwen F. Cancer incidence in the Netherlands in 1989 and 1990: First results of the nationwide Netherlands Cancer Registry. Eur J Cancer. 1995;31:1822-9.
29. De Rubeis V, Bayat S, Griffith LE, Smith BT, Anderson LN. Validity of self-reported recall of anthropometric measures in early life: A systematic review and meta-analysis. Obes Rev. 2019;20:1426-40.
30. Ferrari P, Friedenreich C, Matthews CEJAjoe. The role of measurement error in estimating levels of physical activity. 2007;166:832-40.
31. Patrick DL, Cheadle A, Thompson DC, Diehr P, Koepsell T, Kinne S. The validity of self-reported smoking: a review and meta-analysis. American journal of public health. 1994;84:1086-93.

CHAPTER 9

RISK FACTORS FOR CANCER OF UNKNOWN PRIMARY: A LITERATURE REVIEW

Karlijn E.P.E. Hermans, Fatemeh Kazemzadeh, Caroline Loef, Rob L.H. Jansen, Iris D. Nagtegaal, Piet A. van den Brandt, and Leo J. Schouten

Submitted for publication

Abstract

Background: Cancer of Unknown Primary (CUP) is metastatic cancer with an unidentifiable primary tumour origin during life. Hitherto, it is unclear which risk factors are associated with CUP, yet identifying these factors could reveal whether CUP is a specific entity or a cluster of metastasised cancers from various primary tumour origins.

Objective: To review CUP risk factors.

Data sources: Epidemiological studies on possible CUP risk factors were systematically searched in PubMed on February 1st, 2022.

Study selection: Studies, published before 2022, were included if they were observational human-based, provided relative risk estimates, and investigated possible CUP risk factors.

Data extraction: Relative risk estimates with p-values or 95% confidence intervals were extracted.

Results: A total of 4 case-control and 14 cohort studies were included. There appears to be an increased risk for smoking in relation to CUP. However, limited suggestive evidence was found to link alcohol consumption, diabetes mellitus, and family history of cancer as increased risks for CUP. No conclusive associations could be made for anthropometry, food intake (animal or plant-based), immunity disorders, lifestyle (overall), physical activity, or socioeconomic status and CUP risk. No other CUP risk factors have been studied.

Conclusions and implications: This review highlights smoking, alcohol consumption, diabetes mellitus and family history of cancer as CUP risk factors. Yet, there remains insufficient epidemiological evidence to conclude that CUP has its own specific risk factor profile.

Background

Cancer of unknown primary (CUP) is an aggressive unpredictable metastatic cancer with an unidentifiable primary tumour origin during life [1-4]. The disease predominantly occurs in older individuals with a median age of 60 years [5]. The NICE guideline categorised CUP into 1) malignancy of undefined primary origin (MUO), 2) provisional CUP: metastatic epithelial or neuroendocrine malignancy identified based on microscopical verification, and 3) confirmed CUP: metastatic epithelial or neuroendocrine malignancy identified based on final histology, with no primary site detected despite a selected initial screen of investigations, specialist review, and specialised investigations as appropriate [6, 7]. These categories are useful in clinical settings, but population-based research datasets contain a mixture of CUP cases that do not clearly distinguish provisional from confirmed cases. With this mixture [1, 7, 8], variability in disease registrations and diagnostic workup between countries [8-10], it remains hard to compare CUP occurrence globally and identify time trends.

Cancer risk factors and prevention

The International Agency for Research on Cancer (IARC) Monographs have identified various environmental factors that are carcinogenic hazards to humans, which it continually reviews and updates. These include chemicals, occupational exposures, physical agents, biological agents, and lifestyle factors [11]. Identifying risk factors can guide primary prevention to reduce diseases [12, 13] and for CUP specifically, this is especially important given the bleak prognosis.

Rationale

To the best of our knowledge, one review examined pointers of disease mechanisms associated with CUP [14], yet, in recent years, the epidemiological evidence regarding those pointers has expanded, which is why we provide here a comprehensive review of current CUP risk factors. We have examined risk factors in association with CUP, considering that a risk factor profile for CUP may imply whether CUP is a specific entity or a cluster of metastasised cancers from various primary tumour origins.

Material and methods

The literature search on CUP risk factors (2011-2022) was performed in PubMed on February 1st, 2022 by using the following keywords (MeSH) and free text terms for the exposure groups: alcohol consumption; anthropometry (body mass index, waist circumference, body constitution and waist-hip ratio); diabetes mellitus (DM); drinks (coffee, caffeine, tea); family history of cancer (FHC) (medical history taking, genetic predisposition to disease); foods (vegetables, fruits, meats, fish products, dairy products, milk, soy milk, eggs, soy foods, soybeans, bread, whole grains, cereal, nuts and seeds); physical activity (exercise, sedentary behaviour); smoking (smoking and tobacco smoke pollution); socioeconomic status (SES) (social conditions, income, poverty, socioeconomic factors, employment, unemployment, work, occupations, education, educational status, health, health insurance, health education, health promotion, health behaviour); racial groups and ethnicity; radiation exposure and environmental pollutants (carcinogens); hormonal factors (estrogens, progesterone, testosterone and oral hormonal contraceptives); and reproductive factors (maternal age, menarche, menopause, post-menopausal hormone replacement therapy, parity), in relation to the outcome: neoplasms of unknown primary, also referred to as cancer of unknown primary.

Studies were included if they were observational (e.g., cohort and case-control) human-based, provided risk estimates with p-values or 95% confidence intervals, and/or if they had data on at least one of the abovementioned exposure groups. No language restrictions were used. The reference lists of the included articles were checked for potentially relevant studies. Data were extracted for general characteristics and exposure estimates. Due to variability between the studies concerning the study design, different exposures (including differences in exposure measurement), and differences in confounder adjustment, it was not possible to conduct a pooled meta-analysis. Therefore, the existing epidemiological evidence was compared and described as a comprehensive discussion on CUP risk factors. All studies were evaluated against the World Cancer Research Fund's (WCRF) criteria as epidemiological evidence for cancer prevention, which ranges from convincing to limited-no conclusion. Its criteria are derived from the Bradford Hill criteria which consider the strength of association, temporality, consistency, biological plausibility, dose-response relationship, and experimental evidence [15].

One researcher (K.H.) screened abstracts and eligible full texts, and uncertainties were discussed with a second researcher (L.S.). The reference lists of included articles were checked for additional studies.

Results

The PubMed search yielded 878 articles, 18 articles of which were deemed eligible for inclusion (Figure 1). Overall, seven research teams had examined CUP risk factors in European, American, and Australian populations, representing 4 case-control and 14 cohort studies (Table 1A and 1B). Record linkage methods for exposure and follow-up measurements were applied through country-specific cancer, pathology, and healthcare registers. The search revealed studies on alcohol consumption, anthropometry, DM, FHC, food intake (animal and plant-based), immunity disorders, lifestyle (overall), physical activity, smoking, and SES in relation to CUP risk. (Supplementary Tables A-K). No studies had examined the association between drinks, racial groups and ethnicity, radiation exposure and environmental pollutants, hormonal factors, or reproductive factors, and CUP risk.

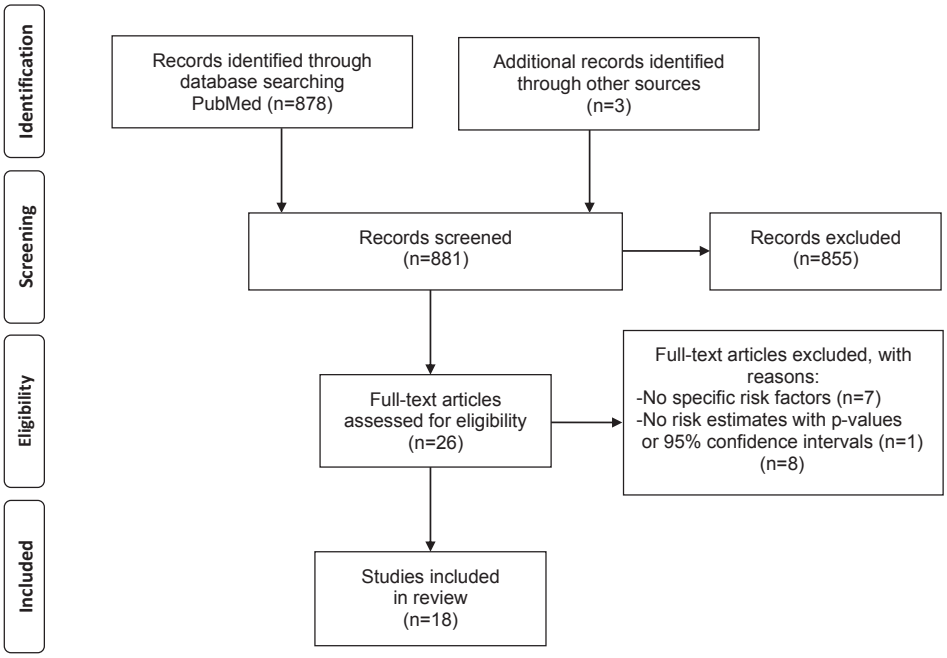


Figure 1 Flowchart of included and excluded studies on whom the review on Cancer of Unknown Primary risk factors is based

Evaluation of results

Based on the grading criteria in relation to CUP risk, convincing – strong evidence was found for smoking, whereas limited to suggestive evidence was seen for alcohol consumption, DM, and FHC, and limited - no conclusive evidence for anthropometry, food intake (animal or plant-based), immunity disorders, lifestyle (overall), physical activity, or SES (Table 2).

Smoking

Four studies explored the association between smoking and CUP risk (Figure 2 & Supplementary Table A). All studies reported statistically significantly increased associations for smoking status in relation to CUP [9, 16-18]. Kaaks et al. & Vajdic et al. also observed an even higher CUP risk among participants who smoked the highest number of cigarettes per day (26+ and 20+, respectively) compared to never smokers. Similarly, Hermans et al. observed a statistically significant association for smoking frequency which became higher with an increasing number of cigarettes smoked compared to never smokers. They also found smoking duration ≥40 years (multivariable adjusted HR: 1.45, 95% CI: 1.09-1.94, *P*trend = 0.02), and smoking cessation (current smokers) associated with increased CUP risk (multivariable adjusted HR: 1.67, 95% CI: 1.37-2.03, *P*trend <0.001) compared to never smokers [17]. Although the strength of the associations varies between these studies, they all point to a positive association between smoking and CUP risk, particularly in the highest exposure categories.

Alcohol consumption

Alcohol consumption related to CUP risk was investigated in three studies (Figure 3 & Supplementary Table B). Kaaks et al. & Hermans et al. reported increased risks for participants in the highest exposure categories of alcohol consumption >60g and ≥30g in relation to CUP compared to 0-12g and abstainers, respectively [9, 17], whereas Vajdic et al. observed no associations between alcohol consumption and CUP risk compared to non-consumers [18]. Despite the different consumption categories and confounder adjustments there is a suggestive relationship between alcohol consumption and CUP risk.

Table 1A. General characteristics of case-control studies on risk factors and cancer of unknown primary.

Reference	Country	Record linkage	CUP-definition	Assessed exposures
Hemminki et al., 2011,2012 (22, 23)	Sweden	Swedish Family-Cancer Database, MigMed2 dataset at the Center for Primary Health Care Research, Lund University, Malmö	ICD: 7, not further specified	Family history of cancer
Hemminki et al., 2014 (16)	Sweden	Swedish prospective biobanks (Umea Medical Biobank and the Malmo Diet and Cancer Study and the Prevention Study)	ICD: 7, 9, 10, not further specified	BMI, smoking
Hemminki et al., 2016 (19)	Sweden	Swedish healthcare registers: national Hospital Discharge Register (diagnoses 1997-2010), national Outpatient Registry (2001-2010), and the Primary Health Care Registry in Stockholm County (2001-2007). Patients were linked to the Swedish Cancer Registry	CUP diagnoses obtained from Swedish Cancer Registry	Diabetes mellitus

Notes:
Abbreviations: BMI: body mass index, FU: follow-up, N.A.: not applicable

Table 1B. General characteristics of cohort studies on risk factors and cancer of unknown primary.

Reference	Country	Study name	Follow-up	Record linkage	CUP-definition	Assessed exposures
Hemminki et al., 2015 (31)	Sweden	N.A.	1964-2012	Swedish Hospital Discharge Register, the KriMed database at Center for Primary Health Care Research, Malmö, Lund University	ICD: 7 code 199, ICD: 9 code 1990-1991	Autoimmune diseases
Kaaks et al., 2014 (9)	Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom	The European Prospective Investigation into Cancer and Nutrition Cohort	1992-2000	Linkages with cancer and pathology registries. Health insurance records (France), direct contacts with participants and their next of kin (Germany and Greece); disease occurrence in the latter countries were systematically verified against clinical and pathological records	ICD-O-2: C809: unknown primary site	Alcohol consumption, anthropometry (BMI and waist circumference), levels of education, smoking

Reference	Country	Study name	Follow-up	Record linkage	CUP-definition	Assessed exposures
Samadder et al., 2016 (24)	Unites States of America	Utah Population Database from the Genealogical Society of Utah	1980-2010	Electronic records for patients with CUP derived from the Utah Cancer Registry	ICD-O-3: C80.9: unknown primary site Histology subtypes: adenocarcinoma (M8140-8389), squamous cell carcinoma (M8050-8089), carcinoma not otherwise specified (M8010-8049), and neuroendocrine (M8240, 8241, 8243-8246, 8249)	Family history of cancer
Urban et al., 2016 (29)	Unites States of America	Surveillance, Epidemiology and End Results (SEER)-files	SEER 9: 1973-1991 SEER 13: 1992-1999 SEER 17: 2000-2008	SEER-program of the National Cancer Institute in the United States.	ICD-O-3: C80.9: unknown primary site Histology subtypes: adenocarcinoma (8140–8389); squamous cell carcinoma (8050–8089); carcinoma not otherwise specified (NOS; 8010–8049) and neuroendocrine (8240, 8241, 8243-8246 and 8249)	SES
Vajdic et al., 2019, 2019 (18, 20)	Australia	The Sax Institute's 45 and Up Study	1994-2012	NSW Cancer Registry (1994-2012), NSW Admitted Patients Data Collection (2001-2015), NSW Emergency Department Data Collection (2005-2016), and NSW Registry of Births, Deaths, and Marriages (2006-2016)	ICD-O-3: C80 (unknown primary site), C76 (other and ill-defined sites), C26 (other and ill-defined digestive organs) or C39 (other and ill-defined sites within respiratory system and intrathoracic organs)	Age, alcohol consumption, BMI, meat consumption, physical activity, SES: educational attainment; employment; household income; private health insurance; and residential location, sex, smoking, vegetable and fruit consumption, self-reported health conditions, self-reported family history of cancer, and hospital-recorded health conditions

Reference	Country	Study name	Follow-up	Record linkage	CUP-definition	Assessed exposures
Pavlidis et al., 2020 (30)	Unites States of America	Surveillance, Epidemiology and End Results (SEER)-files	SEER 9: 1973-1991 SEER 13: 1992-1999 SEER 17: 2000-2008	SEER-program of the National Cancer Institute in the United States.	ICD-O-3: C80.9: unknown primary site Histology subtypes: adenocarcinoma (M8140–8239, 8241-8245, 8247, 8248, 8350-8389); squamous cell carcinoma (M8050–8089); carcinoma with neuroendocrine differentiation (M8240, 8246, 8249)	Ethnicity, residential location, SES
Hermans et al., 2020-2022 (17, 21, 26, 27, 28, 32) & Grewcock et al., 2021 (25)	The Netherlands	The Netherlands Cohort Study	1986-2006	Record linkage with the Netherlands Cancer Registry, Dutch Pathology Registry, and Registration of Municipal Administration	ICD-O-3: C80.9: unknown primary site Histology subtypes: M8000-8570	Alcohol consumption, anthropometry (BMI and clothing size), diabetes mellitus, foods (animal and plant), overall lifestyle, physical activity, smoking & Family history of cancer

Notes:
Abbreviations: BMI: body mass index, FU: follow-up, N.A.: not applicable, NSW: New South Wales, SES: socioeconomic status

Table 2. Exposure evaluation according to the grading criteria as evidence for cancer prevention as reported by the World Cancer Research Fund applied to cancer of unknown primary risk

Exposure	Number of studies & study design	Evaluation of results	Category of evidence *
Alcohol consumption	3 Cohort (9, 17, 18)	- Evidence from at least two independent cohort studies - The direction of effect is generally consistent though some unexplained heterogeneity may be present - Evidence for biological plausibility	Limited – suggestive
Anthropometry	1 Case-control (16) 3 Cohort (9, 18, 26)	- Evidence is so limited that no firm conclusion can be made	Limited – no conclusion
Diabetes mellitus	1 Case-control (19) 2 Cohort (20, 21)	- Evidence from at least two independent cohort studies - The direction of effect is generally consistent though some unexplained heterogeneity may be present - Evidence for biological plausibility	Limited – suggestive
Family history of cancer	2 Case-control (22, 23) 3 Cohort (20, 24, 25)	- Evidence from at least two independent cohort studies - The direction of effect is generally consistent though some unexplained heterogeneity may be present - Evidence for biological plausibility	Limited – suggestive
Foods – animal based	2 Cohort (18, 27)	- Evidence is so limited that no firm conclusion can be made	Limited – no conclusion
Foods – plant based	2 Cohort (18, 28)	- Evidence is so limited that no firm conclusion can be made	Limited – no conclusion
Immunity – autoimmune diseases	1 Case-control (31)	- Evidence is so limited that no firm conclusion can be made	Limited – no conclusion
Lifestyle – overall	1 Cohort (32)	- Evidence is so limited that no firm conclusion can be made	Limited – no conclusion
Physical activity	2 Cohort (18, 26)	- Evidence is so limited that no firm conclusion can be made	Limited – no conclusion

Exposure	Number of studies & study design	Evaluation of results	Category of evidence *
Smoking	1 Case-control (16) 3 Cohort (9, 17, 18)	- Evidence from more than one study type - Evidence from at least two independent cohort studies - No substantial unexplained heterogeneity within or between study types or in different populations relating to the presence or absence of an association, or direction of effect - Good-quality studies to exclude with confidence the possibility that observed association results from random or systematic error, including confounding, measurement error and selection bias - Presence of a plausible biological gradient (dose-response) in the association. Such a gradient need not be linear or even in the same direction across the different levels of exposure, so long as this can be explained plausibly	Convincing – strong evidence
Socioeconomic status	3 Cohort (18, 29, 30)	- Evidence is so limited that no firm conclusion can be made	Limited – no conclusion

Notes
*Category of evidence can be distinguished into the following subgroups as issued by the World Cancer Research Fund: 'convincing', 'probable', 'limited-suggestive', 'limited-no conclusion', and 'substantial effect on risk unlikely'.

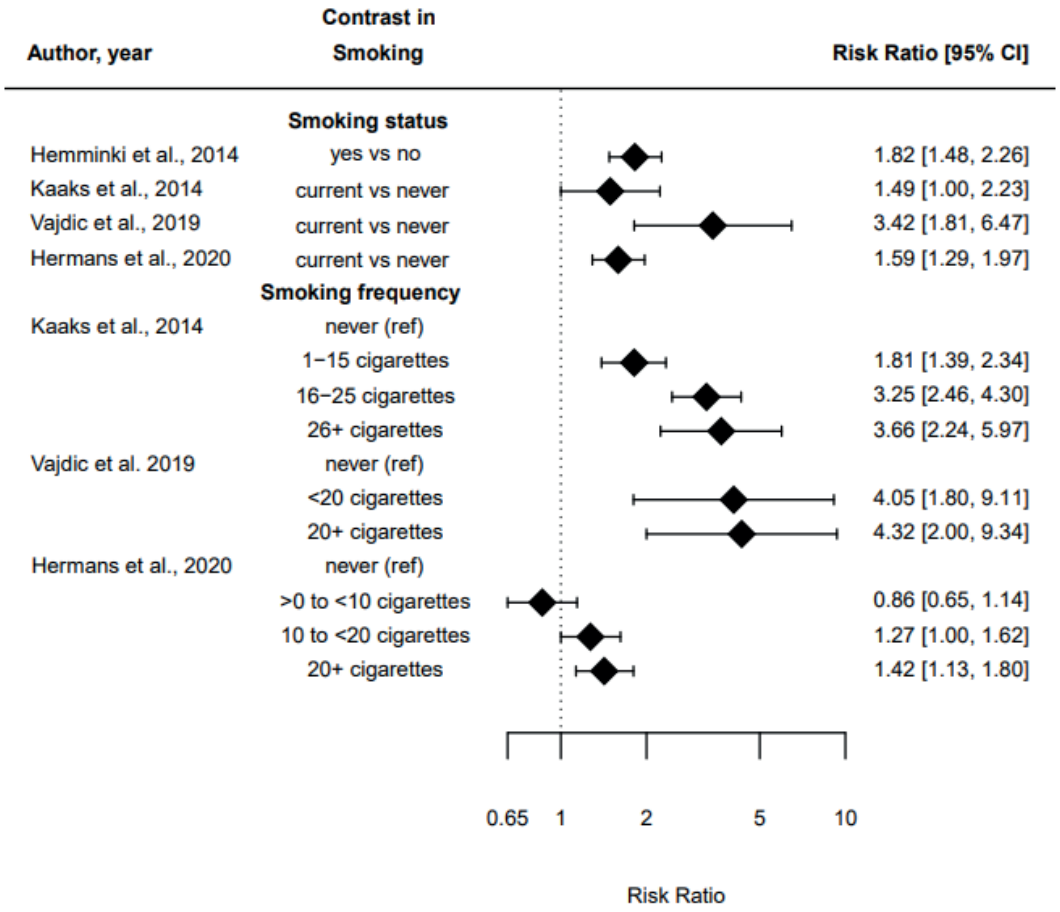


Figure 2 Risk ratios of smoking in relation to Cancer of Unknown Primary risk

Supplementary Table A. Results of epidemiological studies on smoking and cancer of unknown primary risk.

Smoking			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Hemminki et al., 2014	Cases: 290 participants smoked, 173 participants did not smoke Controls: 1140 smoked, 1212 participants did not smoke	**Smoking status** Smoking versus no smoking 1.82 (1.48-2.26), *p*-value: $2.8x10^{-8}$	-
Kaaks et al., 2014	Cohort: 521,448 Cases: 619	**Smoking intensity** Never smoked: referent Current smokers, 1-15 cigarettes: 1.81 (1.39-2.34) Current smokers, 16-25 cigarettes: 3.25 (2.46-4.30) Current smokers, 26+ cigarettes: 3.66 (2.24-5.97) Former smokers, quit ≤10 y: 1.34 (0.99-1.80) Former smokers, quit >10 y: 1.08 (0.86-1.36) Current smokers, pipe or cigar: 1.49 (1.00-2.23)	Levels of education, BMI, waist circumference, and average lifetime alcohol consumption
Vajdic et al., 2019	Cohort:266,933 Cases: 327 General cohort population controls: 981	**Smoking status** Regular tobacco smoking (age-sex adjustment) General cohort population controls Never: referent Former: 2.03 (1.44-2.86) Current: 4.19 (2.33-7.55) Current, <20/day 4.05 (1.80-9.11) Current, ≥20/day 4.32 (2.00-9.34)	Age and sex
		Regular tobacco smoking (multivariable adjustment) General cohort population controls Never: referent Former: 1.95 (1.33-2.86) Current: 3.42 (1.81-6.47)	Age, educational attainment, smoking history, self-rated health, self-reported anxiety, self-reported diabetes, and history of cancer at baseline

Smoking

Reference	Exposure and *N**	Risk estimates (95% CI)	Adjusted for
Hermans et al., 2020	Cohort: 120,852 Subcohort: 4,288 Cases: 963	**Cigarette smoking status** Never smokers: referent Ex-smokers: 1.19 (0.97-1.47) Current smokers: 1.59 (1.29-1.97) *p*-trend: <.001	Age at baseline, sex, alcohol consumption, cigarette smoking frequency, and cigarette smoking duration
		Cigarette smoking frequency Never smokers: referent >0 to <10: 0.86 (0.65-1.14) 10 to <20: 1.27 (1.00-1.62) ≥20: 1.42 (1.13-1.80) *p*-trend: 0.003	Age at baseline, sex, alcohol consumption, current cigarette smoking status, and cigarette smoking duration
		Cigarette smoking duration Never smokers: referent >0 to <20: 0.95 (0.71-1.27) 20 to <40: 1.07 (0.86-1.33) ≥40: 1.45 (1.09-1.94) *p*-trend: 0.02	Age at baseline, sex, alcohol consumption, current cigarette smoking status, cigarette smoking frequency
		Time since cigarette smoking cessation Never smokers: referent Stopped ≥20 y: 0.91 (0.67-1.23) Stopped 10 to <20 y: 1.06 (0.81-1.38) Stopped >0 to <10 y: 1.26 (0.99-1.62) Current smokers: 1.67 (1.37-2.03) *p*-trend: <.001	Age at baseline, sex, alcohol consumption, number of cigarette pack-years

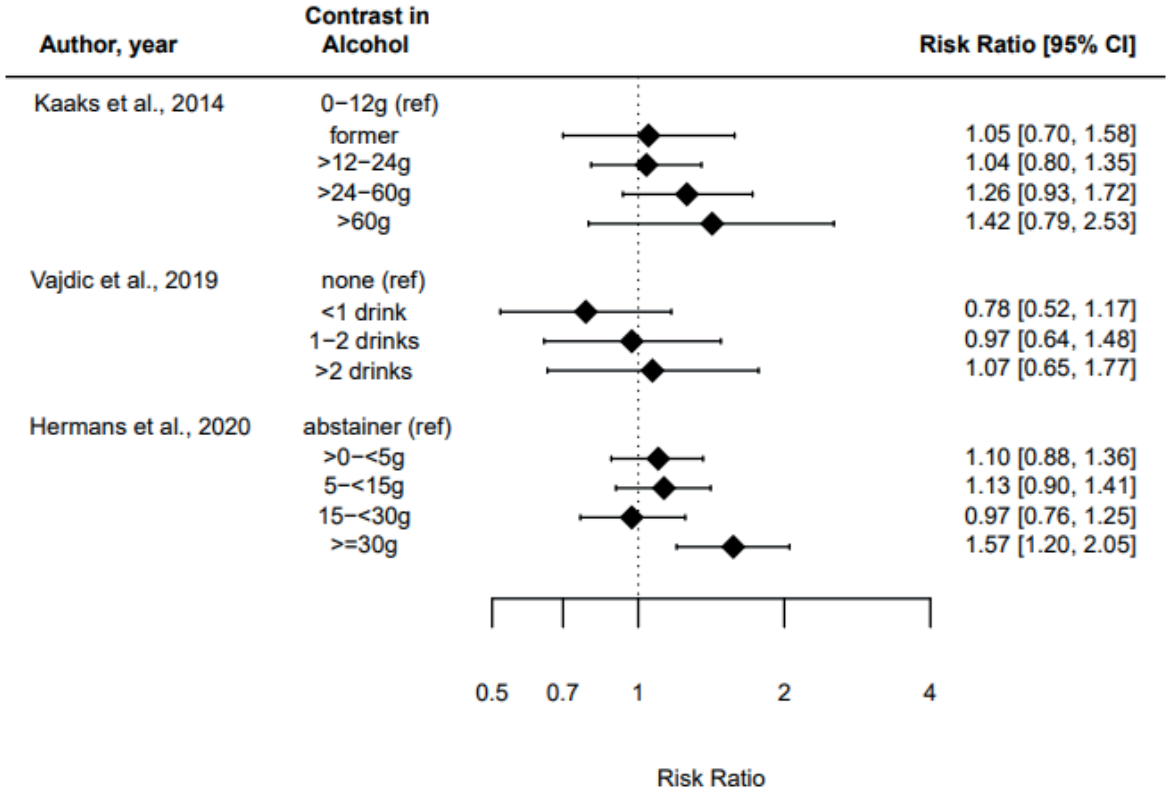


Figure 3 Risk ratios of alcohol consumption in relation to Cancer of Unknown Primary risk

Supplementary Table B. Results of epidemiological studies on alcohol consumption and cancer of unknown primary risk.

Alcohol consumption			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Kaaks et al., 2014	Cohort: 521,448 Cases: 510	**Daily alcohol consumption** Former: 1.05 (0.70-1.58) 0 to 12: referent >12 to 24: 1.04 (0.80-1.35) >24 to 60: 1.26 (0.93-1.72) >60: 1.42 (0.79-2.53) *p*-trend: 0.15	Levels of education, BMI, waist circumference, and smoking intensity
Vajdic et al., 2019	Cohort:266,933 Cases: 327 General cohort population controls: 981	**Daily alcohol consumption** General cohort population controls None referent <1 drink: 0.78 (0.52-1.17) 1-2 drinks: 0.97 (0.64-1.48) >2 drinks: 1.07 (0.65-1.77)	Age and sex
Hermans et al., 2020	Cohort: 120,852 Subcohort: 4,288 Cases: 963	**Daily alcohol consumption** Abstainers: referent >0 to <5: 1.10 (0.88-1.36) 5 to <15: 1.13 (0.90-1.41) 15 to <30: 0.97 (0.76-1.25) ≥30: 1.57 (1.20-2.05) *p*-trend: 0.02	Age at baseline, sex, current cigarette smoking status, cigarette smoking frequency, and cigarette smoking duration

Diabetes mellitus

The association between diabetes mellitus and CUP risk was investigated in three studies (Figure 4 & Supplementary Table C). Hemminki et al., found that participants with Type 1 (T1DM) and Type 2 DM (T2DM) (with or without insulin treatment) had a statistically significantly increased CUP risk compared to participants without DM [19]. Similarly, Vajdic et al. also found a statistically significant relationship between DM and increased CUP risk compared to participants without DM [20]. Lastly, Hermans et al. observed a non-significant association between T2DM and increased CUP risk compared to participants with no DM [21]. Overall, there appears to be a suggestive association between DM and increased CUP risk. Though its strength might be affected due to inability of confounder adjustment.

Family history of cancer

Five studies reported on the association between family history of cancer and CUP risk (Figure 5 & Supplementary Table D). Hemminki et al. found statistically significantly increased CUP risks in siblings alone, while no associations were found between FHC and CUP risk in parents alone [22, 23]. In a follow-up study, Hemminki et al. reported a statistically significantly increased CUP risk in first degree relatives [22]. Similarly, Samadder et al. reported a statistically significant association between family history of cancer and CUP risk in first-degree relatives, but, no associations in second-degree relatives or first cousins [24]. Vajdic et al. reported no association between FHC and CUP risk [20]. Grewcock et al. observed a non-significant increased CUP risk for FHC in siblings only. No associations were found between FHC in parents only in relation to CUP risk [25]. Therefore, there seems to be a suggestive association between FHC and CUP risk. Both Hemminki et al. and Grewcock et al. suggest an association between FHC in siblings only concerning CUP risk, but confounder adjustment was not conducted in the study by Hemminki et al. It is possible that the findings observed for siblings result from a shared environment, which is less likely between the parents and the index-case.

Anthropometry

Four studies investigated the association between anthropometry and CUP risk (Supplementary Table E). Hemminki et al. compared CUP patients with a BMI ≥20 to CUP patients with a BMI <20 (case-control), and found a decreased CUP risk,

albeit not statistically significant [16]. Kaaks et al. found no associations between BMI and CUP risk, but when comparing the highest quartile to the lowest, they did observe that participants with an increasing waist circumference were at an increased CUP risk (multivariable adjusted HR: 1.29, 95% CI: 1.02-1.65, *P*trend = 0.01), which suggests a potential link with abdominal fat [9]. Vajdic et al. noted that obese participants had a non-significant increased CUP risk compared to normal weight participants (age-sex adjusted OR: 1.37, 95% CI: 0.87-2.13) [18]. Hermans et al. explored the association by investigating height (sex-stratified), BMI at baseline, BMI at age 20 years, change in BMI since age 20 years, and clothing size as a proxy for waist circumference (trouser size for men, skirt size for women), but even after multivariable adjustments found no associations between these variables in relation to CUP risk [26].

Foods (animal-based)

Vajdic et al. and Hermans et al. investigated consuming animal foods and CUP risk (Supplementary Table F). Neither study found any association in respect to red meat consumption. However, Vajdic et al. found an inverse association between processed meat consumption and CUP risk (age-sex adjusted OR: 1.28, 95% CI: 0.82-1.99) compared to consumers <3 meat per week [18], while Hermans et al. found a statistically significantly increased CUP risk for participants with the highest consumption (Q4) of processed meats compared to the lowest consumption (Q1) categories (multivariable adjusted HR: 1.40, 95% CI: 1.12-1.75, *P*trend = 0.006) [27].

Foods (plant-based)

Two studies investigated plant foods consumption in relation to CUP risk. Vajdic et al. reported that participants with an intake of ≥5 vegetables per day, or an intake of ≥2 fruits per day, had a non-significant decreased CUP risk (age-sex adjusted OR: 0.79, 95% CI: 0.57-1.10 & OR: 0.73, 95% CI: 0.53-1.00, respectively) compared to consuming <5 vegetables per day, and <2 fruits per day [18]. Hermans et al. studied vegetable and fruit consumption as a group, and as individual components for vegetables, legumes, and fruits, but found no associations between any (Q4) of the plant food exposures in relation to CUP risk compared to the lowest intake (Q1) categories (multivariable adjusted HR: 0.97, 95% CI: 0.78-1.20, *P*trend = 0.63) [28] (Supplementary Table G).

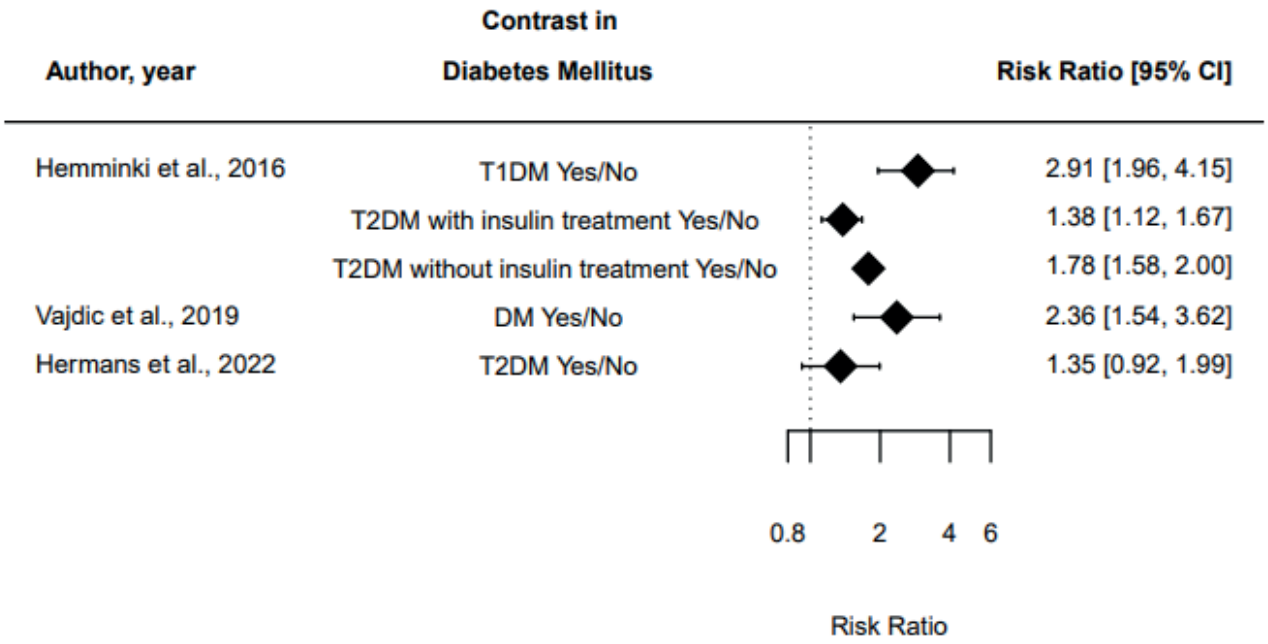


Figure 4 Risk ratios of diabetes mellitus in relation to Cancer of Unknown Primary risk

Supplementary Table C. Results of epidemiological studies on diabetes mellitus and cancer of unknown primary risk.

Diabetes mellitus			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Hemminki et al., 2016	T1DM Cases: 32,600 T2DM Cases:178,000	**T1DM status** SIR: 2.91 (1.96-4.15) **T2DM status with insulin treatment** SIR: 1.38 (1.12-1.67) **T2DM status without insulin treatment** SIR: 1.78 (1.58-2.00)	-
Vajdic et al., 2019	Cohort:266,933 Cases: 327 General cohort population controls: 981	**Diabetes status** General cohort population controls Yes/no: 2.36 (1.54-3.62)	Age and sex
Hermans et al., 2022	Cohort: 120,852 Subcohort: 4,288 Cases: 963	**T2DM status** Yes/no: 1.35 (0.92-1.99)	Age at baseline, sex, alcohol consumption, current cigarette smoking status, cigarette smoking frequency, and cigarette smoking duration

Notes:
Abbreviations: DM: diabetes mellitus (T1DM: Type 1, T2DM: Type 2)

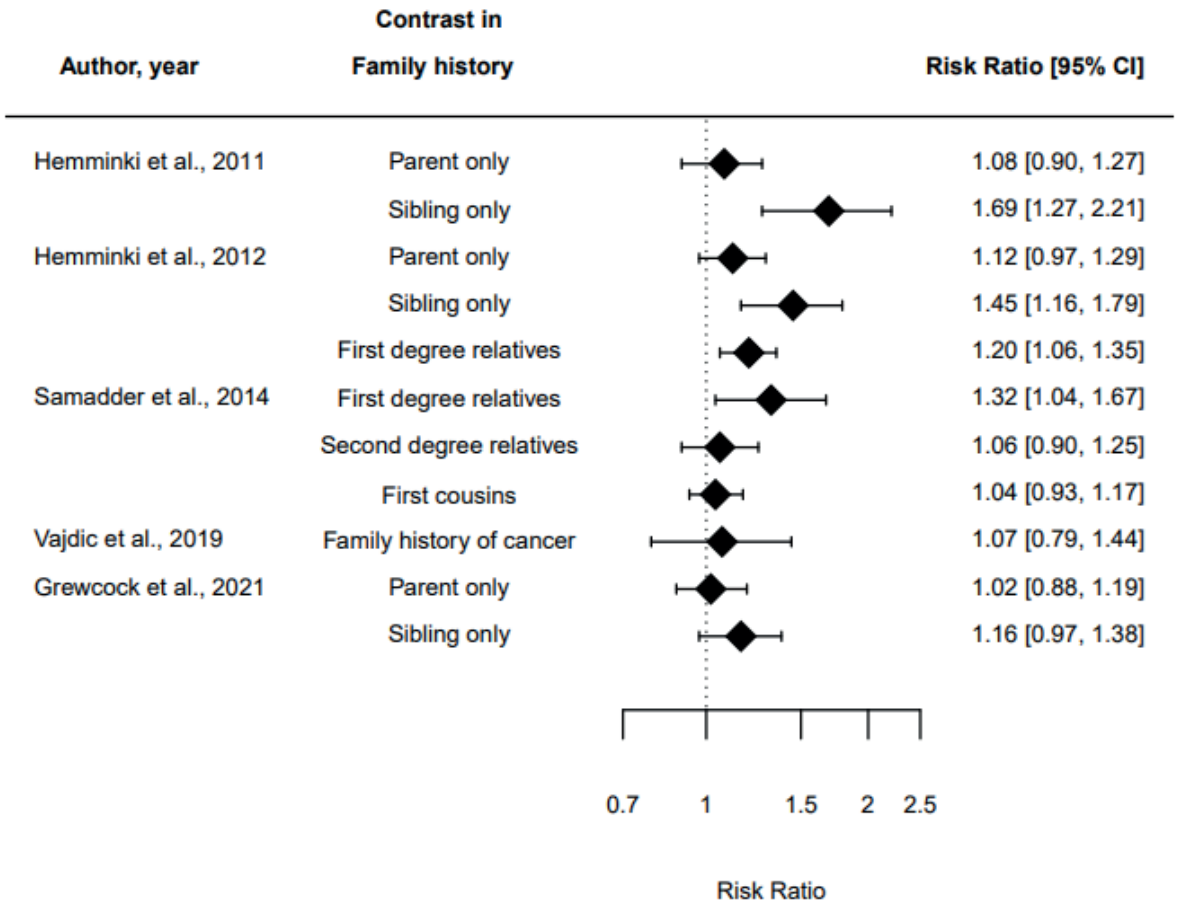


Figure 5 Risk ratios of family history of cancer in relation to Cancer of Unknown Primary risk

Supplementary Table D. Results of epidemiological studies on family history of cancer and cancer of unknown primary risk.

Family history of cancer			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Hemminki et al., 2011	Offspring individuals: 9,558,512 Cases: 35,168	**Family history of cancer - Parent only** SIR: 1.08 (0.90-1.27) **Family history of cancer - Sibling only** SIR: 1.69 (1.27-2.21)	-
Hemminki et al., 2012	Offspring individuals: 9,171 with CUP; 5,506 (60%) had a first-degree relative with any cancer Cases: 56,049	**Family history of cancer - Parent only** SIR: 1.12 (0.97-1.29) **Family history of cancer - Sibling only** SIR: 1.45 (1.16-1.79) **Family history of cancer - First degree relatives** SIR: 1.20 (1.06-1.35)	-
Samadder et al., 2015	Cases: 4,160 Controls, non-CUP: 52,036 Controls, cancer-free: 51,053	**Family history of cancer - First-degree relatives** cancer-free controls: 1.32 (1.04-1.67) **Family history of cancer - Second-degree relatives** cancer-free controls: 1.06 (0.90-1.25) **Family history of cancer - First cousins** cancer-free controls: 1.04 (0.93-1.17)	-
Vajdic et al., 2019	Cohort:266,933 Cases: 327 General cohort population controls: 981	**Family history of cancer** General cohort population controls Yes/no: 1.07 (0.79-1.44)	Age and sex
Grewcock et al., 2021	Cohort: 120,852 Subcohort: 4,288 Cases: 963	**Family history of cancer – Parent only** HR: 1.02 (0.88-1.19) **Family history of cancer – Sibling only** HR: 1.16 (0.97-1.38)	Age at baseline, sex, alcohol consumption, current cigarette smoking status, cigarette smoking frequency, and cigarette smoking duration

Supplementary Table E. Results of epidemiological studies on anthropometry and cancer of unknown primary risk.

Anthropometry			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Hemminki et al., 2014	Cases: 418 BMI ≥20, 29 BMI <20 Controls: 2203 BMI ≥20, 113 BMI <20	**BMI** BMI ≥20 versus BMI <20 (All CUP) 0.77 (0.50-1.18), *p*-value: 0.23	-
Kaaks et al., 2014	Cohort: 521,448 Cases BMI: 634 Cases waist circumference: 600	**BMI** Quartile 1: referent Quartile 2: 0.92 (0.73-1.16) Quartile 3: 0.98 (0.78-1.23) Quartile 4: 1.06 (0.84-1.33) *p*-trend: 0.29 **Waist circumference** Quartile 1: referent Quartile 2: 0.91 (0.71-1.16) Quartile 3: 1.02 (0.80-1.30) Quartile 4: 1.29 (1.02-1.65) *p*-trend: 0.01	Smoking intensity, average lifetime alcohol consumption, and levels of education
Vajdic et al., 2019	Cohort:266,933 Cases: 327 General cohort population controls: 981	**BMI** General cohort population controls Underweight: 2.13 (0.52-8.61) Normal weight: referent Overweight: 0.89 (0.61-1.31) Obese: 1.37 (0.87-2.13)	Age and sex
Hermans et al., 2020	Cohort: 120,852 Subcohort: 4,099 Cases: 926	**Height (men only)** <170: referent 170-<175: 0.90 (0.65-1.24) 175-<180: 0.85 (0.61-1.19) 180-<185: 0.88 (0.61-1.27) ≥185: 0.91 (0.59-1.41) *p*-trend: 0.67 **Height (women only)** <160: referent 160-<165: 0.74 (0.51-1.06) 165-<170: 0.77 (0.54-1.09) 170-<175: 1.03 (0.69-1.51) ≥175: 0.99 (0.59-1.67) *p*-trend: 0.62	Age at baseline, sex, alcohol consumption, current cigarette smoking status, cigarette smoking frequency, and cigarette smoking duration, and weight

Anthropometry

Reference	Exposure and *N**	Risk estimates (95% CI)	Adjusted for
Hermans et al., 2020	Cohort: 120,852 Subcohort: 4,099 Cases: 926	**BMI at baseline** <20: 0.91 (0.58-1.42) 20-<25: referent 25-<30: 0.90 (0.77-1.06) ≥30: 1.11 (0.81-1.52) *p*-trend: 0.77	Age at baseline, sex, alcohol consumption, current cigarette smoking status, cigarette smoking frequency, and cigarette smoking duration, and non-occupational physical activity
		BMI at age 20 years <20: 0.89 (0.71-1.13) 20-<21.5: referent 21.5-<23: 0.93 (0.74-1.17) 23-<25: 0.96 (0.75-1.23) ≥25: 1.06 (0.85-1.32) *p*-trend: 0.25	Age at baseline, sex, alcohol consumption, current cigarette smoking status, cigarette smoking frequency, and cigarette smoking duration, and non-occupational physical activity
		Change in BMI since age 20 years <0: 1.24 (0.97-1.61) 0-<4: referent 4-<8: 0.82 (0.68-0.99) ≥8: 1.04 (0.86-1.26)	Age at baseline, sex, alcohol consumption, current cigarette smoking status, cigarette smoking frequency, and cigarette smoking duration, non-occupational physical activity, and BMI at age 20 years
		Clothing size (proxy for waist circumference) Clothing size men (trouser size) <50: 0.94 (0.68-1.30) 50-51: referent 52-53: 0.82 (0.62-1.08) 54-55: 0.76 (0.54-1.05) >56: 0.93 (0.68-1.28) *p*-trend: 0.46 Clothing size women (skirt size) <40: 0.91 (0.63-1.30) 42: referent 44: 0.99 (0.72-1.36) 46-48: 0.87 (0.62-1.22) >50: 1.51 (0.91-2.52) *p*-trend: 0.55	Age at baseline, sex, alcohol consumption, current cigarette smoking status, cigarette smoking frequency, and cigarette smoking duration, and non-occupational physical activity

Notes:
Abbreviations: BMI: body mass index

Supplementary Table F. Results of epidemiological studies on foods (*animal-based*) and cancer of unknown primary risk.

Animal foods			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Vajdic et al., 2019	Cohort:266,933 Cases: 327 General cohort population controls: 981	**Red meat consumption** General cohort population controls ≥3 meat/week: 0.95 (0.69-1.32) **Processed meat consumption** General cohort population controls ≥3 meat/week: 1.28 (0.82-1.99)	Age and sex
Hermans et al., 2021	Cohort: 120,852 Subcohort: 4,111 Cases: 899	**Red meat (overall)** Q1: referent Q2: 1.11 (0.89-1.37) Q3: 1.21 (0.98-1.49) Q4: 1.04 (0.83-1.30) *p*-trend: 0.31 **Poultry** C1: referent C2: 1.13 (0.91-1.41) C3: 1.10 (0.88-1.37) C4: 0.97 (0.79-1.21) *p*-trend: 0.28 **Processed meat** Q1: referent Q2: 1.07 (0.86-1.33) Q3: 1.14 (0.91-1.42) Q4: 1.40 (1.12-1.75) *p*-trend: 0.006 **Fish** Q1: referent Q2: 1.30 (1.04-1.61) Q3: 1.23 (1.00-1.51) Q4: 1.25 (0.99-1.57) *p*-trend: 0.29	Age at baseline, sex, current cigarette smoking status, cigarette smoking frequency, cigarette smoking duration, and total energy intake

Physical activity

Two studies have reported on the relationship between physical activity and CUP risk (Supplementary Table H). Vajdic et al. found that participants who were physically active for >150 minutes per week (total and moderate-vigorous physical activity) had a statistically significant decreased CUP risk (age-sex adjusted OR: 0.63, 95% CI: 0.44-0.88) compared to participants who were physically active for <150 minutes per week. They also found that physically active participants, >2 times per week, had an even lower CUP risk (age-sex adjusted OR: 0.48, 95% CI: 0.26-0.89) compared to <1 times per week [18]. Hermans et al. studied non-occupational physical activity in relation to CUP risk but found no association after multivariable adjustment when comparing participants who were physically active for >90 minutes per day to ≤30 minutes per day [26].

Socioeconomic status

Urban et al. found neither educational level nor poverty to be associated with CUP risk [29] (Supplementary Table I). Vajdic et al. explored components of SES in relation to CUP risk and found participants without school certificate/qualification to be more at risk (multivariable adjusted OR: 1.69, 95% CI: 1.08-2.64) than participants with any school certificate/ qualification. Additionally, disabled/sick participants, or unemployed, had increased CUP risks. Those who held private health insurance had a decreased CUP risk. In terms of income, participants with a lower income, or who did not report their income, had increased CUP risks [18]. In contrast, Pavlidis et al. reported that participants with a high SES had an increased association for CUP risk (RR: 1.90, 95% CI: 1.50-2.60) compared to those with a low SES [30]. Vajdic et al. suggest that a poor SES measured by education, employment, and access to health services, is associated with increased CUP risk, although these findings may differ between populations. Its authors did not report on adjustments for smoking behaviour or alcohol consumption, while both exposures are linked to SES and may thus play an influential role in the association with CUP. In contrast, Pavlidis et al. in their adjusted analysis, found that participants with a higher SES had a higher CUP risk, while in their unadjusted analysis they found a protective risk. Unfortunately, they did not clarify which variables they had adjusted for in the analysis, so it is impossible to rule out potential correlation between variables.

Immunity disorders

One case-control study, by Hemminki et al. investigated whether dysfunctions of the immune system in autoimmune diseases are linked to increased CUP risk (Supplementary Table J). It found an overall increased CUP risk for patients diagnosed with autoimmune diseases (SIR: 1.27, 95% CI: 1.22-1.32) [31]. However, the researchers could not control for smoking, which may have influenced the association.

Lifestyle (overall)

Hermans et al. examined whether adhering to lifestyle recommendations, as issued by the WCRF and American Institute for Cancer Research in 2018 for cancer prevention helps in decreasing CUP risk. Lifestyle was defined as including a healthy weight, physical activity, and the consumption of plant and animal foods, and alcohol. The highest adherence to lifestyle recommendations was significantly associated with a decreased CUP risk in the age-sex adjusted analysis compared to lowest adherence. However, after adjusting for smoking as well the association attenuated (multivariable adjusted HR: 0.87, 95% CI: 0.70-1.08) [32] (Supplementary Table K).

Supplementary Table G. Results of epidemiological studies on foods (*plant-based*) and cancer of unknown primary risk.

Plant foods			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Vajdic et al., 2019	Cohort:266,933 Cases: 327 General cohort population controls: 981	**Vegetable consumption** General cohort population controls ≥5 vegetables/day: 0.79 (0.57-1.10)	Age and sex
		Fruit consumption General cohort population controls ≥2 fruits/day: 0.73 (0.53-1.00)	
Hermans et al., 2021	Cohort: 120,852 Subcohort: 4,005 Cases: 867	**Total vegetables and fruits (combined)** Q1: referent Q2: 1.02 (0.83-1.27) Q3: 0.96 (0.78-1.19) Q4: 0.97 (0.78-1.20) *p*-trend: 0.63	Age at baseline, sex, current cigarette smoking status, cigarette smoking frequency, and cigarette smoking duration
		Total vegetables Q1: referent Q2: 0.94 (0.76-1.17) Q3: 1.04 (0.84.-1.28) Q4: 0.87 (0.69-1.09) *p*-trend: 0.38	Age at baseline, sex, current cigarette smoking status, cigarette smoking frequency, cigarette smoking duration, and total fruit consumption
		Legumes Q1: referent Q2: 1.11 (0.90-1.38) Q3: 1.08 (0.87.-1.35) Q4: 1.21 (0.97-1.52) *p*-trend: 0.14	Age at baseline, sex, current cigarette smoking status, cigarette smoking frequency, cigarette smoking duration, and total vegetable and fruit consumption
		Total fruits Q1: referent Q2: 0.94 (0.76-1.16) Q3: 0.92 (0.74.-1.15) Q4: 0.94 (0.75-1.17) *p*-trend: 0.56	Age at baseline, sex, current cigarette smoking status, cigarette smoking frequency, cigarette smoking duration, and total vegetable consumption

Supplementary Table H. Results of epidemiological studies on physical activity and cancer of unknown primary risk.

Physical activity			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Vajdic et al., 2019	Cohort:266,933 Cases: 327 General cohort population controls: 981	**Total and moderate vigorous physical activity** General cohort population controls >150 min/week: 0.63 (0.44-0.88) **Total physical activity** General cohort population controls <1 times/week: referent 1-2 times/week: 0.95 (0.38-2.33) >2 times/week: 0.48 (0.26-0.89)	Age and sex
Hermans et al., 2020	Cohort: 120,852 Subcohort: 4,099 Cases: 926	**Non-occupational physical activity in minutes per day** ≤30: referent >30-60: 0.91 (0.74-1.12) >60-90: 0.85 (0.67-1.08) >90: 0.97 (0.78-1.20) *p*-trend: 0.84	Age at baseline, sex, alcohol consumption, current cigarette smoking status, cigarette smoking frequency, cigarette smoking duration, and BMI at baseline

Supplementary Table I. Results of epidemiological studies on socioeconomic status and cancer of unknown primary risk.

Socioeconomic status			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Urban et al., 2013	Cohort: >2,800,000 Cases: 51,294	**Educational level** Lowest education: referent Second quartile: 0.96 (0.94-0.99) Third quartile: 0.97 (0.95-1.00) Highest education: 0.95 (0.93-0.98) **Poverty** Affluent: referent Middle: 1.02 (1.00-1.04) Poor: 1.03 (0.98-1.07)	-
Vajdic et al., 2019	Cohort:266,933 Cases: 327 General cohort population controls: 981	**Educational attainment (6-categories)** General cohort population controls No school certificate or qualification: referent School of intermediate certificate: 0.47 (0.29-0.75) Higher school or leaving certificate: 0.49 (0.26-0.94) Trade/apprenticeship: 0.75 (0.44-1.29) Certificate/diploma: 0.50 (0.30-0.85) University degree or higher: 0.51 (0.30-0.86) **Employment** General cohort population controls Full-time, part-time, or self-employed: 0.85 (0.56-1.28) Fully or part retired: 0.85 (0.58-1.27) Disabled/sick: 3.33 (1.67-6.65) Doing unpaid work: 0.63 (0.30-1.34) Studying: 1.50 (0.35-6.40) Looking after home/family: 0.92 (0.52-1.61) Unemployed: 2.71 (1.01-7.31) **Hold private health insurance** General cohort population controls Yes: 0.62 (0.45-0.85)	Age and sex

Socioeconomic status			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Vajdic et al., 2019	Cohort:266,933 Cases: 327 General cohort population controls: 981	**Residential location** Major city: referent Inner regional: 1.01 (0.73-1.41) Outer regional or rural: 1.46 (0.90-2.35) Not reported: 0.59 (0.17-1.98)	Age and sex
		Yearly household income General cohort population controls $70,000 or more: referent $50,000-$69,999: 1.38 (0.67-2.85) $20,000-$49,999: 1.30 (0.73-2.30) <$20,000: 1.91 (1.08-3.37) Not reported: 1.79 (1.00-3.21)	
		Educational attainment (2-categories) General cohort population controls Any school certificate or qualification: referent No school certificate or qualification: 1.69 (1.08-2.64)	Age, educational attainment, smoking history, self-rated health, self-reported anxiety, self-reported diabetes, and history of cancer at baseline
Pavlidis et al., 2020	Cases: 907 adolescents and young adults	**Ethnicity** Black: 1.17 (1.00-1.40) Other: 1.01 (0.80-1.20) White: Reference	**Unclear which variables are adjusted for in the analysis*
		Urban Nonmetropolitan: 0.16 (0.10-0.20) Metropolitan: Reference	
		Socioeconomic status High: 1.90 (1.50-2.60) Low: Reference	

Notes:
Abbreviations: SES: socioeconomic status

Supplementary Table J. Results of epidemiological studies on immunity disorders and cancer of unknown primary risk.

Immunity disorders				
Reference	**Exposure and *N****	**Risk estimates (95% CI)**		**Adjusted for**
Hemminki et al., 2015	Cohort: 789,681 Cases: 2,658	**Autoimmune diseases**		-
		All:	1.27 (1.22-1.32)	
		Addison's disease:	1.77 (1.07-2.78)	
		Celiac disease:	1.55 (1.14-2.05)	
		Crohn's disease:	1.59 (1.37-1.83)	
		Graves'/hyperthyroidism:	1.28 (1.15-1.42)	
		Myasthenia gravis:	1.43 (1.23-1.65)	
		Pemicious anemia:	1.55 (1.29-1.84)	
		Polymyalgia rheumatica:	1.21 (1.05-1.38)	
		Polymyositis//dermatomyositis:	3.51 (2.44-4.89)	
		Primary biliary cirrhosis:	1.81 (1.05-2.91)	
		Psoriasis:	1.15 (1.01-1.30)	
		Rheumatoid arthritis:	1.14 (1.05-1.25)	
		Sjögren's syndrome:	1.55 (1.21-1.96)	
		Systemic lupus erythematosus:	1.57 (1.15-2.09)	
		Systemic sclerosis:	1.60 (1.16-2.14)	
		Ulcerative colitis:	1.54 (1.34-1.77)	

Supplementary Table K. Results of epidemiological studies on lifestyle (overall) and cancer of unknown primary risk.

Lifestyle (overall)			
Reference	**Exposure and *N****	**Risk estimates (95% CI)**	**Adjusted for**
Hermans et al., 2022	Cohort: 120,852 Subcohort: 3,911 Cases: 856	**Overall adherence to lifestyle recommendations on cancer prevention; including healthy weight, physical activity, plant and animal foods, and alcohol consumption** Lowest adherence: referent Medium adherence: 1.08 (0.90-1.31) Highest adherence: 0.87 (0.70-1.08) *p*-trend: 0.99	Age at baseline, sex, current cigarette smoking status, cigarette smoking frequency, cigarette smoking duration, and total energy intake

Discussion

Based on epidemiological evidence from 4 case-control and 14 cohort studies reviewed here, there is an association between smoking and CUP risk, but evidence for alcohol consumption, DM, and FHC is limited suggestive. The evidence does not allow conclusive associations to be made for anthropometry, food intake (animal or plant-based), immunity disorders, lifestyle (overall), physical activity, or SES.

Explanation of findings

Autopsy results from CUP patients indicate that primary tumours tend to originate in the lung(s) (5-35%) or pancreas (15-20%), and less often the liver and bile ducts (10-15%), or colon/rectum (3-8%)[33]. These higher occurrences for the lung and pancreas may be reflective of the associations observed with smoking and alcohol consumption. After all, it is known that smoking is strongly associated with lung cancer through deregulated cells, cytokines, and growth factors, which may elevate epithelial apoptosis resistance and ultimately result in mutations [14, 34]. Higher levels of alcohol consumption may be linked to primary tumours of the mouth, pharynx, larynx, oesophagus, liver, pancreas, breast and colorectum [9, 35], the mechanisms underlying cancer development include DNA, protein, and lipid alterations, or damage by acetaldehyde, as well as the carcinogenic metabolite of ethanol, oxidative stress, and alterations to hormonal regulations [36].

For DM, other mechanisms may play a role as patients with T2DM generally have an impaired immune system [37]. Studies have reported that T2DM is related to various types of cancer [38], which may influence the ability of the immune system to suppress a primary tumour, but that the metastasis escaped immune suppression [1, 19, 39]. Similarly, when studying FHC, the role of genetic or environmental risk factors may also be indicative of a specific cancer type. Participants were found to have an increased CUP risk if they had a FHC including kidney, colorectal, lung, pancreatic, myeloma and non-Hodgkin lymphoma [22-24]. These cancer types may be reflective of the primary tumour origin in the CUP patients.

It remains unclear as to whether CUP is a specific entity, or whether there are specific mechanisms that explain its pattern of metastasis. One of the mechanisms that could explain the absence of indicating a primary tumour origin is, as briefly indicated above, that the immune system was able to remove the primary tumour

after metastasis as the primary tumour is recognized, but unable to distinguish features of the metastasis and therefore discard the metastasis in some CUP cases [14, 40]. Studies on CUP immune profiling have shown similar immune profiles compared to immune therapy responsive malignancies [41-43]. Some differences in immune responses to foreign and self-antigens are present throughout life, while others depend on gene expressions and hormone status. These differences may be influenced by gender, early environmental exposures, race, and, for example, systemic inflammatory autoimmune diseases [44-46]. In addition, the genes involved in the immune system are under constant evolutionary pressure due to pathogens, environmental conditions, and the relocation of populations [44, 45, 47]. The findings here indicate associations with smoking, alcohol consumption, DM, and FHC in relation to CUP risk, and these risk factors are all known to negatively affect the immune system's ability to intercept malignant cell development [1, 19, 22, 23, 39, 48, 49]. Similar findings have been found in a study that evaluated immunity disorders in relation to CUP risk [31]. Due to the immune system's (in) ability to intercept, and the association found between immunity disorders and CUP occurrence, one could speculate that the immune system and CUP incidence are correlated.

Implications

This literature review examined various factors and showed that smoking, alcohol consumption, DM, and FHC appear to be associated with CUP risk. The heterogeneous nature of CUP as well as the lack of a specific aetiology suggest that CUP is not a specific entity. Indeed, it is more likely that CUP is a cluster of metastasised cancers, which would explain the variation in both aetiology and immunology.

Future CUP studies

A novel approach to study specific aspects of a disease is computational pathology. This approach enables scientists to use sources of information, including patients' histology data, to extract patterns of cancer. Studies have used artificial intelligence based on both molecular information as well as routine histology slides to investigate the feasibility of predicting the tumour of origin in CUP patients [50]. This procedure could potentially reduce the extensive diagnostic work-ups that patients undergo. Therefore, future studies into the epidemiological risk factors of

CUP besides studying metastatic patterns of cancers with known primaries to learn about the progression model of cancers and combining them with computational pathology predictions for CUP could accelerate the diagnostic process and identify the tumour of origin so as to help personalize therapies [51-54].

Validity and methodological considerations of the epidemiological findings

CUP risk factors have rarely been studied, most probably due to the lack of a consistent disease definition and because of a general lack of awareness. This dearth of research already makes comparisons hard, but that task is made even harder because those studies that have been done have tended to apply different study designs, used different definitions of the outcome measure, used different exposure assessments, and differences in availability of confounder data. The lack of confounder data collection restricts confounder adjustments in the analyses, and consequently, associations may have been under- or overestimated.

Conclusions

This review has highlighted the influence of a healthy lifestyle on CUP risk, and shown that while there does appear to be an increased risk for smoking, there is only limited suggestive evidence for alcohol consumption, DM, and FHC. No conclusive associations can be made for anthropometry, food intake (animal or plant-based), immunity disorders, lifestyle (overall), physical activity, or SES and CUP risk. Consequently, there is insufficient epidemiological evidence to conclude that CUP has its own specific risk factor profile.

References

1. Loffler H, Kramer A. Introduction. In: Loffler H, Kramer A, editors. Cancer of Unknown Primary Springer International Publishing Switzerland; 2016.
2. Pavlidis N, Khaled H, Gaafar R. A mini review on cancer of unknown primary site: A clinical puzzle for the oncologists. J Adv Res. 2015;6:375-82.
3. [3] Pavlidis N. Cancer of unknown primary: biological and clinical characteristics. Annals of oncology : official journal of the European Society for Medical Oncology. 2003;14 Suppl 3:iii11-8.
4. Lee MS, Sanoff HK. Cancer of unknown primary. BMJ. 2020;371.
5. Pavlidis N. Forty years experience of treating cancer of unknown primary. Acta oncologica (Stockholm, Sweden). 2007;46:592-601.
6. National Institute for Health and Care Excellence. Metastastic malignant disease of unknown primary origin. Diagnosis and management of metastatic disease of unknown primary origin. NICE clinical guideline 104. 2010.
7. Brewster DH, Lang J, Bhatti LA, Thomson CS, Oien KA. Descriptive epidemiology of cancer of unknown primary site in Scotland, 1961-2010. Cancer Epidemiol. 2014;38:227-34.
8. Schroten-Loef C, Verhoeven R, de Hingh I, van de Wouw A, van Laarhoven H, Lemmens V. Unknown primary carcinoma in the Netherlands: decrease in incidence and survival times remain poor between 2000 and 2012. Eur J Cancer. 2018;101:77-86.
9. Kaaks R, Sookthai D, Hemminki K, Kramer A, Boeing H, Wirfalt E, et al. Risk factors for cancers of unknown primary site: Results from the prospective EPIC cohort. Int J Cancer. 2014;135:2475-81.
10. Iris van der Strate; Fatemeh Kazemzadeh; Iris D. Nagtegaal; Debbie Robbrecht; Yes vd Wouw; Catarina S. Padilla; Saskia Duijts; Manel Esteller FAGNP. Global Consensus on the Initial Diagnostic Workup of Cancer of Unknown Primary Submitted, personal communication. 2022.
11. International Agency for Research on Cancer. Monographs on the identification of carcinogenic hazards to humans. IARC Monographs - general information2022.
12. Dovjak M, Kukec A. Creating Healthy and Sustainable Buildings: An Assessment of Health Risk Factors: Springer Nature; 2019.
13. Wild CP, Bucher JR, De Jong BW, Dillner J, Von Gertten C, Groopman JD, et al. Translational cancer research: balancing prevention and treatment to combat cancer globally. J Natl Cancer Inst. 2015;107:1-5.
14. Hemminki K. Epidemiology, Risk Factors, and Survival in CUP: Pointers to Disease Mechanisms. Cancer of Unknown Primary: Springer; 2016. p. 5-25.
15. World Cancer Research Fund/American Institute for Cancer Research. Judging the evidence. Continuous Update Project2018.
16. Hemminki K, Chen B, Melander O, Manjer J, Hallmans G, Hemminki A. Smoking and body mass index as risk factors for subtypes of cancer of unknown primary. Int J Cancer. 2015;136:246-7.
17. Hermans KE, van den Brandt PA, Loef C, Jansen RL, Schouten LJ. Alcohol consumption, cigarette smoking and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study. Int J Cancer. 2021.

18. Vajdic C, Perez-Concha O, Dobbins T, Ward RL, Schaffer AL, van Leeuwen M, et al. Demographic, social and lifestyle risk factors for cancer registry-notified cancer of unknown primary site (CUP). Cancer Epidemiol. 2019;60:156-61.
19. Hemminki K, Försti A, Sundquist K, Li X. Cancer of unknown primary is associated with diabetes. Eur J Cancer Prev. 2016;25:246-51.
20. Vajdic C, Perez-Concha O, Rhee JJ, Dobbins T, Ward RL, Schaffer AL, et al. Health-related predictors of cancer registry-notified cancer of unknown primary site (CUP). Cancer Epidemiol. 2019;61:1-7.
21. Hermans KE, Haque S, van den Brandt PA, Loef C, Jansen RL, Schouten LJ. Type 2 diabetes mellitus and Cancer of Unknown Primary risk: results from the Netherlands Cohort Study. Eur J Cancer Prev. 2022.
22. Hemminki K, Bevier M, Sundquist J, Hemminki A. Cancer of unknown primary (CUP): does cause of death and family history implicate hidden phenotypically changed primaries? Annals of oncology : official journal of the European Society for Medical Oncology. 2012;23:2720-4.
23. Hemminki K, Ji J, Sundquist J, Shu X. Familial risks in cancer of unknown primary: tracking the primary sites. J Clin Oncol. 2011;29:435-40.
24. Samadder NJ, Smith KR, Hanson H, Pimentel R, Wong J, Boucher K, et al. Familial risk in patients with carcinoma of unknown primary. JAMA Oncology. 2016;2:340-6.
25. Grewcock AL, Hermans KE, Weijenberg MP, van den Brandt PA, Loef C, Jansen RL, et al. Family history of cancer in first degree relatives and risk of cancer of unknown primary. Eur J Cancer Care. 2021.
26. Hermans KE, van den Brandt PA, Loef C, Jansen RL, Schouten LJ. Anthropometry, physical activity and cancer of unknown primary (CUP) risk: Results from the Netherlands cohort study. Cancer Epidemiol. 2020;69:101836.
27. Hermans KEPE, Van den Brandt PA, Loef C, Jansen RLH, LJ. S. Meat consumption and cancer of unknown primary (CUP) risk: results from The Netherlands cohort study on diet and cancer. Eur J Nutr. 2021.
28. Hermans K, Van den Brandt P, Loef C, Jansen R, Schouten L. Vegetable and fruit consumption and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study on Diet and Cancer. BMC Cancer. 2022;22.
29. Urban D, Rao A, Bressel M, Lawrence YR, Mileshkin L. Cancer of unknown primary: a population-based analysis of temporal change and socioeconomic disparities. Br J Cancer. 2013;109:1318-24.
30. Pavlidis N, Rassy E, Smith-Gagen J. Cancer of unknown primary: Incidence rates, risk factors and survival among adolescents and young adults. Int J Cancer. 2020;146:1490-8.
31. Hemminki K, Sundquist K, Sundquist J, Ji J. Risk of cancer of unknown primary after hospitalization for autoimmune diseases. Int J Cancer. 2015;137:2885-95.
32. Hermans KEPE, van den Brandt PA, Loef C, Jansen RLH, Schouten LJ. Adherence to the World Cancer Research Fund and the American Institute for Cancer Research lifestyle recommendations for cancer prevention and Cancer of Unknown Primary risk. Clinical Nutrition. 2022;41:526-35.
33. Neben K, Hübner G, Folprecht G, Jäger D, Krämer A. Metastases in the absence of a primary tumor: advances in the diagnosis and treatment of CUP syndrome. J Deutsches Ärzteblatt International. 2008;105:733.

34. Walser T, Cui X, Yanagawa J, Lee JM, Heinrich E, Lee G, et al. Smoking and lung cancer: the role of inflammation. J Proceedings of the American Thoracic Society. 2008;5:811-5.
35. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. Continuous Update Project Expert Report 20182018.
36. Rumgay H, Shield K, Charvat H, Ferrari P, Sornpaisarn B, Obot I, et al. Global burden of cancer in 2020 attributable to alcohol consumption: a population-based study. The Lancet Oncology. 2021;22:1071-80.
37. Berbudi A, Rahmadika N, Tjahjadi AI, Ruslami R. Type 2 Diabetes and its Impact on the Immune System. Current diabetes reviews. 2020;16:442-9.
38. Giovannucci E, Harlan DM, Archer MC, Bergenstal RM, Gapstur SM, Habel LA, et al. Diabetes and cancer: a consensus report. Diabetes care. 2010;33:1674-85.
39. Pavlidis N, Pentheroudakis G. Cancer of unknown primary site. Lancet. 2012;379:1428-35.
40. Krämer A, Löffler H. Cancer of Unknown Primary: Springer International Publishing; 2015.
41. Hainsworth JD, Greco FA. Cancer of Unknown Primary Site: New Treatment Paradigms in the Era of Precision Medicine. 2018:20-5.
42. Gatalica Z, Xiu J, Swensen J, Vranic S. Comprehensive analysis of cancers of unknown primary for the biomarkers of response to immune checkpoint blockade therapy. Eur J Cancer. 2018;94:179-86.
43. Haratani K, Hayashi H, Takahama T, Nakamura Y, Tomida S, Yoshida T, et al. Clinical and immune profiling for cancer of unknown primary site. Journal for immunotherapy of cancer. 2019;7:251.
44. Klein SL, Flanagan KL. Sex differences in immune responses. Nature reviews Immunology. 2016;16:626-38.
45. Roved J, Westerdahl H, Hasselquist D. Sex differences in immune responses: Hormonal effects, antagonistic selection, and evolutionary consequences. Hormones and behavior. 2017;88:95-105.
46. Özdemir BC, Dotto GP. Racial Differences in Cancer Susceptibility and Survival: More Than the Color of the Skin? Trends in cancer. 2017;3:181-97.
47. Quach H, Rotival M, Pothlichet J, Loh YE, Dannemann M, Zidane N, et al. Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. Cell. 2016;167:643-56.e17.
48. Romeo J, Wärnberg J, Nova E, Díaz LE, Gómez-Martinez S, Marcos A. Moderate alcohol consumption and the immune system: A review. British Journal of Nutrition. 2007;98:S111-S5.
49. Qiu F, Liang CL, Liu H, Zeng YQ, Hou S, Huang S, et al. Impacts of cigarette smoking on immune responsiveness: Up and down or upside down? Oncotarget. 2017;8:268-84.
50. Lu MY, Chen TY, Williamson DFK, Zhao M, Shady M, Lipkova J, et al. AI-based pathology predicts origins for cancers of unknown primary. Nature. 2021;594:106-10.
51. de Ridder J, de Wilt JH, Simmer F, Overbeek L, Lemmens V, Nagtegaal I. Incidence and origin of histologically confirmed liver metastases: an explorative case-study of 23,154 patients. Oncotarget. 2016;7:55368-76.

52. Hugen N, Sloot YJE, Netea-Maier RT, van de Water C, Smit JWA, Nagtegaal ID, et al. Divergent Metastatic Patterns Between Subtypes of Thyroid Carcinoma Results From the Nationwide Dutch Pathology Registry. The Journal of clinical endocrinology and metabolism. 2020;105:e299-306.
53. Verstegen MH, Harker M, van de Water C, van Dieren J, Hugen N, Nagtegaal ID, et al. Metastatic pattern in esophageal and gastric cancer: Influenced by site and histology. World journal of gastroenterology. 2020;26:6037-46.
54. Knijn N, van Erning FN, Overbeek LI, Punt CJ, Lemmens VE, Hugen N, et al. Limited effect of lymph node status on the metastatic pattern in colorectal cancer. Oncotarget. 2016;7:31699-707.

10

CHAPTER 10

GENERAL DISCUSSION

Summary of the main findings

This thesis aimed to investigate the association between individual lifestyle components: 1) alcohol consumption, cigarette smoking, anthropometry, physical activity, vegetable and fruit consumption, meat consumption, family history of cancer, and diabetes mellitus in relation to CUP risk, and lifestyle as an overall component by studying: 2) whether adherence to the WRCF/AICR lifestyle recommendations for cancer prevention is associated with CUP risk. Finally, we discussed the findings of previous epidemiological studies in combination with those observed in the NLCS in an up-to-date comprehensive review. In this overview, we discuss the most important findings from the NLCS-studies and those of the comprehensive review.

Alcohol consumption (dose-response relationship) and cigarette smoking appeared to be associated with an statistically significantly increased CUP risk (see Chapter 2) (1). We observed no associations for anthropometry or physical activity in relation to CUP development (see Chapter 3) (2). Overall, vegetable and fruit consumption did not appear to be associated with CUP risk (see Chapter 4) (3). Beef and processed meat consumption were found to be statistically significantly associated with increased CUP risk, whereas no associations were observed for red meat (overall), poultry, or fish consumption (see Chapter 5) (4). Family history of cancer was not found to be an independent risk factor in relation to CUP risk (see Chapter 6) (5). A non-significant positive association was observed between T2DM status and CUP risk (see Chapter 7) (6). In our adherence study, we observed that participants with the highest adherence to the recommendations had a statistically significant decreased CUP risk in the age and sex adjusted analysis (see Chapter 8) (7). In our comprehensive review, smoking was found to be an established increased risk factor, while limited suggestive evidence was observed for alcohol consumption, diabetes mellitus, and family history of cancer, and no conclusive associations were found for anthropometry, intake of foods (animal or plant-based), immunity disorders, lifestyle (overall), physical activity, or socioeconomic status in relation to CUP risk (see Chapter 9) (8).

Methodological considerations and future recommendations

To adequately investigate the epidemiology and aetiology of CUP, it is important to utilise a consistent disease definition. Until now, there is no international consensus on a specific CUP definition. The difference in the case-mix makes it very difficult to compare CUP occurrence and risk factors on a global scale as well as in time trends.

To study risk factors for CUP we utilised data from the NLCS. The NLCS was initiated in 1986 and participants were followed-up for 20.3 years (September 1986-December 2006). In this timespan, the definition and criteria to register CUP cases were revised and updated by the NCR. In addition, the quality of NCR data depends on the availability of clinical data. For example, many CUP cases were not extensively investigated due to the age, comorbidities, performance status, localisation of the metastasis, and the patient's decision. Our NLCS-dataset, therefore, consists of a varying case-mix based on definitions and criteria. It is also important to realise that both the diagnostic and pathologic accuracy of examinations improved since the start of the cohort study. Due to these improvements, it is very likely that participants with current CUP diagnoses underwent more extensive examinations to identify the primary tumour localisation than those available in the past.

Furthermore, some exposure trends have changed since the start of the NLCS. For example, prevention measures such as reduced availability, increased costs via taxation, health warnings, and marketing bans were proved to be useful for lowering the prevalence of exposure to tobacco smoking (9, 10). Although the exposure to tobacco smoking decreased, alcohol consumption patterns continue to change internationally, while the prevalence of obesity and overweight, and physical inactivity has increased drastically (11, 12). This increased prevalence of obesity and overweight is worrying as it increases risk of chronic disease morbidity (including disability, depression, Type 2 diabetes mellitus, cardiovascular disease, and certain cancers), and disease mortality (12), whereas physical inactivity is known to negatively affect non-communicable diseases and mental health (11).

The NLCS measurement was conducted at baseline in 1986, the stability of the dietary habits was determined from five annually repeated measurements of nutrient intake (13). It should be acknowledged that, currently, there is a greater variability and accessibility of foods and beverages compared to that in 1986, consequently, dietary behaviour may have subsequently changed. For example,

some products that were seasonal in 1986, are now available all year round. To acquire more knowledge on possible prevalence changes of the abovementioned exposures, it would be advised to repeat the execution of repeated measurements. Due to the increased awareness and adherence to an overall healthy lifestyle through diet, physical, mental and social wellbeing, it may be that the exposure to certain modifiable risk factors changes in time. It is, therefore, advisable to continue the examination of exposures in relation to CUP risk on both the national and international scale, and update its findings.

Another methodological consideration is the heterogeneity of the obtained data with respect to subtypes of CUP. As earlier indicated, the quality of the dataset depends on the registrations as recorded by the cancer registry. In the NLCS-dataset it was beneficial to have obtained the data from the cancer registry combined with supplementation of pathology excerpts. This supplementation allowed us to complete missing information and to verify the basis for diagnosis. In general, analyses of CUP subgroups with respect to histology, number of metastases, localisation of the metastasis, and survival duration may be reflective of clinically and aetiologically relevant classifications of primary tumour origins. For example, for metastases with known primary tumour origins it was observed that the median survival for adenocarcinoma with metastases in the large intestine and rectum is approximately eight months, while it is about two months for primary cancer of the liver (14). We have tried to conduct heterogeneity tests for the CUP subgroups in the NLCS-dataset. However, due to small number of participants per category within these subgroups, and large number of missing data, there was insufficient power to detect a clinical effect. For future studies, it may be very interesting to explore the association in a larger context to acquire more insight into primary tumour origins.

Within the NLCS-analyses, we have deliberately decided to solely investigate epithelial CUP malignancies, thereby excluding sarcoma, lymphoma, mesothelioma, and melanoma. The reasoning for excluding these malignancies derives from their infrequent occurrence which results in too few numbers for reliable statistical analyses, as well as their dissimilar aetiology compared to tumours with known primaries. Precisely, for the reason that those malignancies have a different aetiology, it would be very interesting to assess whether there are differences between the epithelial and non-epithelial malignancies when it

comes to risk factors and primary tumour origins. Future studies are, therefore, also encouraged to assess associations in non-epithelial malignancies.

Another methodological consideration that we would like to emphasize is the statistical analysis of adherence to the World Cancer Research Fund (WCRF) and the American Institute for Cancer Research (AICR) lifestyle recommendations for cancer prevention in relation to CUP risk. The WCRF/AICR cancer prevention recommendations represent healthy lifestyle choices with respect to a healthy weight, physical activity, the consumption of wholegrains, vegetables, fruits, and beans, the limited consumption of fast foods, red- and processed meats, the limited consumption of sugar sweetened drinks and alcoholic beverages, supplement use, breastfeeding, and after a cancer diagnosis (15). In addition, they indicate to avoid tobacco use in any form, but the avoidance of tobacco use is not included as an individual measure in the cancer prevention recommendations. For CUP specifically, our NLCS-study (1) and other epidemiological studies have demonstrated that smoking is an important risk factor (16-18). In our adherence study, we have measured the overall adherence association while adjusting for age and sex, and additional adjustments for smoking (status, frequency, duration) and total energy intake. We checked whether there were substantial differences between the age and sex adjusted analysis and the multivariable analysis with and without adjustment for total energy intake. The comparison revealed that in the age and sex adjusted analysis; the highest adherence to the WCRF/AICR cancer prevention recommendations was statistically significantly associated with a decreased CUP risk, whereas the association attenuated in the multivariable adjusted analysis, for which the additional analysis revealed that this particular attenuation derived from the correction for smoking variables alone (7). Due to the findings of our NLCS-studies and the other epidemiological studies, we would like to emphasize that future studies should carefully examine lifestyle risk factors for CUP and the possible confounding effects stemming from exposure to smoking. The possibility of confounder adjustment for smoking behaviour with respect to status, frequency and/or duration is highly needed to correctly examine associations for CUP, as the effects of smoking are likely to influence the estimation, particularly when examining overall lifestyle components in relation to CUP development.

Conclusion

We have investigated the association between individual lifestyle components: 1) alcohol consumption, cigarette smoking, anthropometry, physical activity, vegetable and fruit consumption, meat consumption, family history of cancer, and diabetes mellitus in relation to CUP risk, and lifestyle as an overall component by studying: 2) whether adherence to the WRCF/AICR lifestyle recommendations for cancer prevention is associated with CUP risk. The NLCS findings indicate statistically significant associations between alcohol consumption, cigarette smoking, and meat consumption (beef and processed meats) and increased CUP risk. We also found positive but non-significant associations for family history of cancer and diabetes mellitus in relation to CUP development, while we found no associations for anthropometry, physical activity, or vegetable and/or fruit consumption. Our adherence study showed that participants with the highest adherence to the WCRF/AICR lifestyle recommendations on cancer prevention have a statistically significant decreased CUP risk in the age and sex adjusted analysis. Findings of our comprehensive review revealed smoking, alcohol consumption, diabetes mellitus, and family history of cancer to be the most important associations in relation to CUP development. In conclusion, adhering to a healthy lifestyle appears to be beneficial in the prevention of CUP, which is of great importance as the disease is associated with a bleak prognosis.

References

1. Hermans KE, van den Brandt PA, Loef C, Jansen RL, Schouten LJ. Alcohol consumption, cigarette smoking and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study. Int J Cancer. 2021.
2. Hermans KE, van den Brandt PA, Loef C, Jansen RL, Schouten LJ. Anthropometry, physical activity and cancer of unknown primary (CUP) risk: Results from the Netherlands cohort study. Cancer Epidemiol. 2020;69:101836.
3. Hermans K, Van den Brandt P, Loef C, Jansen R, Schouten L. Vegetable and fruit consumption and Cancer of Unknown Primary (CUP) risk: results from the Netherlands Cohort Study on Diet and Cancer. BMC Cancer. 2022;22.
4. Hermans KEPE, van den Brandt PA, Loef C, Jansen RLH, Schouten LJ. Meat consumption and cancer of unknown primary (CUP) risk: results from The Netherlands cohort study on diet and cancer. European Journal of Nutrition. 2021.
5. Grewcock AL, Hermans KE, Weijenberg MP, van den Brandt PA, Loef C, Jansen RL, et al. Family history of cancer in first degree relatives and risk of cancer of unknown primary. Eur J Cancer Care. 2021.
6. Hermans KE, Haque S, van den Brandt PA, Loef C, Jansen RL, Schouten LJ. Type 2 diabetes mellitus and Cancer of Unknown Primary risk: results from the Netherlands Cohort Study. Eur J Cancer Prev. 2022.
7. Hermans KEPE, van den Brandt PA, Loef C, Jansen RLH, Schouten LJ. Adherence to the World Cancer Research Fund and the American Institute for Cancer Research lifestyle recommendations for cancer prevention and Cancer of Unknown Primary risk. Clinical Nutrition. 2022;41(2):526-35.
8. Hermans KEPE, Kazemzadeh F, Loef C, Jansen RLH, Nagtegaal ID, Van den Brandt PA, et al. Risk factors for Cancer of Unknown Primary: a literature review. Submitted, personal communication. 2022.
9. Rumgay H, Shield K, Charvat H, Ferrari P, Sornpaisarn B, Obot I, et al. Global burden of cancer in 2020 attributable to alcohol consumption: a population-based study. The Lancet Oncology. 2021;22(8):1071-80.
10. Flor LS, Reitsma MB, Gupta V, Ng M, Gakidou E. The effects of tobacco control policies on global smoking prevalence. Nature Medicine. 2021;27(2):239-43.
11. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. The Lancet Global health. 2018;6(10):e1077-e86.
12. Hruby A, Hu FB. The Epidemiology of Obesity: A Big Picture. PharmacoEconomics. 2015;33(7):673-89.
13. Goldbohm RA, van 't Veer P, van den Brandt PA, van 't Hof MA, Brants HA, Sturmans F, et al. Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. Eur J Clin Nutr. 1995;49(6):420-9.
14. Hemminki K. Epidemiology, Risk Factors, and Survival in CUP: Pointers to Disease Mechanisms. Cancer of Unknown Primary: Springer; 2016. p. 5-25.
15. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. 2018.

16. Hemminki K, Chen B, Melander O, Manjer J, Hallmans G, Hemminki A. Smoking and body mass index as risk factors for subtypes of cancer of unknown primary. Int J Cancer. 2015;136(1):246-7.
17. Kaaks R, Sookthai D, Hemminki K, Kramer A, Boeing H, Wirfalt E, et al. Risk factors for cancers of unknown primary site: Results from the prospective EPIC cohort. Int J Cancer. 2014;135(10):2475-81.
18. Vajdic C, Perez-Concha O, Dobbins T, Ward RL, Schaffer AL, van Leeuwen M, et al. Demographic, social and lifestyle risk factors for cancer registry-notified cancer of unknown primary site (CUP). Cancer Epidemiol. 2019;60:156-61.

10

ADDENDUM

SUMMARY

SAMENVATTING (NL)

IMPACT

DANKWOORD

CURRICULUM VITAE

LIST OF PUBLICATIONS

Summary

Cancer of Unknown Primary (CUP) is a metastatic cancer with no identifiable primary tumour origin. In most cancer cases, there is a clear onset of the primary tumour and its progression, but sometimes metastases are the first symptom while the primary tumour cannot be found despite the completion of initial diagnostic workup and histological and/or cytological verification.

This thesis aimed to investigate the association between individual lifestyle components: 1) alcohol consumption, cigarette smoking, anthropometry, physical activity, vegetable and fruit consumption, meat consumption, family history of cancer, and diabetes mellitus in relation to CUP risk, and lifestyle as an overall component by studying: 2) whether adherence to the WRCF/AICR lifestyle recommendations for cancer prevention is associated with CUP risk. Finally, we discussed the findings of previous epidemiological studies in combination with those observed in the NLCS in an up-to-date comprehensive review. In this overview, we present the most important findings from the NLCS-studies and those of the comprehensive review.

To study risk factors associated with CUP, we utilised data from the Netherlands Cohort Study on diet and cancer (NLCS). This prospective cohort includes a study population of 120,852 participants (58,279 men and 62,573 women) aged 55-69 years at baseline in 1986. Participants originated from 204 Dutch municipal population registries. All participants completed a mailed, self-administered questionnaire on dietary habits and other cancer risk factors at baseline in 1986. The questionnaire was evaluated for its validity and reproducibility. Incident CUP cases were identified through annual record linkage of the full cohort with the Netherlands Cancer Registry (NCR) and the Dutch Pathology Registry (PALGA). Participants were followed up for 20.3 years (from 17 September 1986 until 31 December 2006).

Alcohol consumption and cigarette smoking appeared to be associated with an increased CUP risk (see Chapter 2). For alcohol consumption, we observed a dose-response relationship, which reflects participants with the highest ethanol intake levels to have a higher CUP risk. The sex-stratified analysis revealed that men with the highest intake level of ethanol had an even higher CUP risk, while the association slightly attenuated in women. Cigarette smoking status, cigarette smoking frequency, cigarette smoking duration, and time since cigarette smoking cessation were all found to be statistically significantly associated with increased

CUP risk. For these smoking variables, we also found dose-response relationships which indicate that the more a participant is exposed to smoking, the higher the CUP risk gets.

Our findings indicated that neither anthropometry nor physical activity are associated with CUP risk (see Chapter 3). We investigated various aspects of anthropometry including height, BMI at baseline, BMI at age 20 years, change in BMI since age 20 years, and clothing size (trouser size for men/skirt size for women), we found none of these variables to be associated with CUP risk. Physical activity was measured through a non-occupational physical activity value, for which no association with CUP risk was found.

Overall vegetable and fruit consumption did not appear to be associated with CUP risk (see Chapter 4). We have evaluated the consumption of combined groups of vegetables and fruits as well as individual items. We observed no associations between total vegetable and fruit consumption, total vegetables, cooked vegetables, raw vegetables, legumes, brassica vegetables, allium vegetables, cooked leafy vegetables, total fruits, or citrus fruits in relation to CUP risk. The consumption of raw leafy vegetables appeared to decrease CUP risk, although this may be a chance finding. Individual vegetable and fruit items did neither appear to be associated with CUP risk.

Beef and processed meat consumption were found to be statistically significantly associated with increased CUP risk (see Chapter 5). The sex-stratified analysis indicated that the association with beef and processed meat consumption and CUP risk became stronger in women and remained statistically significant. For men, the association between processed meat consumption and CUP risk was no longer statistically significant, albeit the association was still positive. No associations were observed between red meat (overall), poultry, or fish consumption and CUP risk.

Family history of cancer was not found to be an independent risk factor within our study (see Chapter 6). We did observe a moderately increased CUP risk in participants who reported a sibling with cancer compared to those who did not, and we found a slightly increased CUP risk in participants with a family history of cancer in a sister. No association was seen for parents or participants with a brother with family history of cancer. CUP did not appear to be associated with family history of cancer of breast, ovarian, endometrial, bowel, stomach, lung, prostate, bladder, pancreas, head and neck, lymphoma and/or leukaemia. We did

observe a reduced CUP risk in participants who reported a family history of kidney cancer, though this was only based on three cases. This finding may, therefore, be a chance finding.

A non-significant positive association between T2DM status and CUP risk was observed in our study (see Chapter 7). The sex-stratified analysis revealed that the association became stronger in women. Participants who were aged <50 years at diagnosis of T2DM were found to have a statistically significant increased CUP risk, which again became markedly stronger in women alone.

In another study, we investigated whether adherence to the WCRF/AICR lifestyle recommendations on cancer prevention was associated with CUP risk. We examined adherence with respect to body fatness, physical activity, plant foods, meat consumption and alcohol. We observed that participants with the highest adherence to the recommendations had a statistically significant decreased CUP risk in the age and sex adjusted analysis (see Chapter 8). In the multivariable analysis, we observed that the association between adherence to the recommendations and CUP risk was no longer statistically significant after additional adjustments for smoking behaviour. Participants with the highest adherence for the recommendations on meat (red and processed meats) and alcohol consumption were found to have a statistically significantly decreased CUP risk. Adherence to the recommendations with respect to body fatness, physical activity, or intake of plant foods was not associated with CUP risk.

In our comprehensive review, we systematically searched for epidemiological studies on possible CUP risk factors (see Chapter 9). The existing epidemiological evidence describes associations between smoking, family history of cancer, diabetes mellitus, waist circumference, and immunity disorders in relation to CUP risk, whereas weaker associations were found for alcohol consumption, educational attainment, and no associations were found for intake of animal- or plant-based foods. To evaluate the risk factors observed in the NLCS-studies and the existing epidemiological evidence, we utilised the grading criteria as evidence for cancer prevention as reported by the WCRF; ranging from convincing to limited-no conclusion evidence. By applying these grading criteria, smoking appears to be an established increased risk factor for CUP, while limited suggestive evidence was found for alcohol consumption, diabetes mellitus, and family history of cancer, while no conclusive associations were found for anthropometry, intake of foods

(animal or plant-based), immunity disorders, lifestyle (overall), physical activity, or socioeconomic status in relation to CUP risk.

This thesis closes with a summary of the main findings, methodological considerations and future recommendations, and conclusion. Overall, our studies revealed smoking, alcohol consumption, diabetes mellitus, and family history of cancer to be the most important associations in relation to CUP development. In conclusion, adhering to a healthy lifestyle appears to be beneficial in the prevention of CUP, which is of great importance as the disease is associated with a bleak prognosis.

Samenvatting (NL)

Primaire Tumor Onbekend (PTO) is een uitgezaaide kanker zonder identificeerbare primaire tumoroorsprong. In de meeste gevallen van kanker is er een duidelijk begin van de primaire tumor en zijn progressie, maar soms zijn metastasen het eerste symptoom terwijl de primaire tumor niet kan worden gevonden ondanks de voltooiing van initiële diagnostische onderzoeken en histologische en/of cytologische verificatie.

Dit proefschrift had tot doel het verband te onderzoeken tussen individuele leefstijlcomponenten: 1) alcoholconsumptie, het roken van sigaretten, antropometrie, lichamelijke activiteit, groente- en fruitconsumptie, vleesconsumptie, familiegeschiedenis van kanker, en diabetes mellitus in relatie tot PTO-risico, en leefstijl als een algemeen onderdeel door te bestuderen: 2) of naleving van de WRCF/AICR-leefstijlaanbevelingen voor kankerpreventie geassocieerd is met PTO-risico. Ten slotte bespraken we de bevindingen van eerdere epidemiologische onderzoeken in combinatie met die waargenomen in de NLCS in een uitgebreid review. In dit overzicht presenteren we de belangrijkste bevindingen uit de NLCS-onderzoeken en die van het uitgebreide review.

Om risicofactoren geassocieerd met PTO-risico te bestuderen, gebruikten we gegevens van de Nederlandse Cohortstudie naar voeding en kanker (NLCS). Dit prospectieve cohort omvat een studiepopulatie van 120.852 deelnemers (58.279 mannen en 62.573 vrouwen) in de leeftijd van 55-69 jaar in 1986. De deelnemers waren afkomstig uit 204 Nederlandse gemeentelijke bevolkingsregisters. Alle deelnemers vulden een uitgestuurde, zelf ingevulde vragenlijst in over voedingsgewoonten en andere risicofactoren voor kanker in 1986. De vragenlijst werd geëvalueerd op validiteit en reproduceerbaarheid. Incidente PTO-gevallen warden geïdentificeerd door jaarlijkse registratiekoppeling van het volledige cohort met het Integraal Kankercentrum Nederland (IKNL) en de Nederlandse Pathologieregistratie (PALGA). De deelnemers werden 20,3 jaar gevolgd (van 17 september 1986 tot 31 december 2006).

Alcoholconsumptie en het roken van sigaretten bleken geassocieerd te zijn met een verhoogd PTO-risico (zie hoofdstuk 2). Voor alcoholconsumptie hebben we een dosis-responsrelatie waargenomen, die weerspiegelt dat deelnemers met de hoogste ethanolinname een hoger PTO-risico hebben. In de naar geslacht gestratificeerde analyse bleek dat mannen met de hoogste inname van ethanol

een nog hoger PTO-risico hadden, terwijl de associatie bij vrouwen enigszins afnam. Het roken van sigaretten, de frequentie van het roken van sigaretten, de duur van het roken van sigaretten en de tijd sinds het stoppen met roken bleken allemaal statistisch significant geassocieerd te zijn met een verhoogd PTO-risico. Voor deze rookvariabelen vonden we ook dosis-responsrelaties die aangeven dat hoe meer een deelnemer wordt blootgesteld aan roken, hoe hoger het PTO-risico wordt.

Onze bevindingen gaven aan dat zowel antropometrie en lichamelijke activiteit niet geassocieerd zijn met PTO-risico (zie hoofdstuk 3). We hebben verschillende aspecten van antropometrie onderzocht, waaronder lengte, BMI bij baseline, BMI op 20-jarige leefijd, verandering in BMI sinds 20-jarige leeftijd en kledingmaat (broekmaat voor mannen/rokmaat voor vrouwen), géén van deze variabelen waren geassocieerd met PTO-risico. Lichamelijke activiteit werd gemeten door middel van een niet-beroepsmatige fysieke activiteitswaarde, waarvoor geen verband werd gevonden met PTO-risico.

De totale groente- en fruitconsumptie leek niet geassocieerd te zijn met PTO-risico (zie hoofdstuk 4). We hebben de consumptie van gecombineerde groepen groente en fruit en die van individuele items geëvalueerd. We hebben geen verband gevonden tussen totale groente- en fruitconsumptie, totale groenten, gekookte groenten, rauwe groenten, peulvruchten, koolsoorten, alliumgroenten, gekookte bladgroenten, totaal fruit of citrusvruchten in relatie tot PTO-risico. De consumptie van rauwe bladgroenten bleek het PTO-risico te verminderen, hoewel dit een toevalsbevinding kan zijn. Individuele groente- en fruititems leken niet geassocieerd te zijn met CUP-risico.

De consumptie van rundvlees en bewerkt vlees bleek statistisch significant te zijn geassocieerd met een verhoogd PTO-risico (zie hoofdstuk 5). In de naar geslacht gestratificeerde analyse bleek dat de associatie van rundvlees en bewerkt vlees in relatie tot PTO-risico toenam bij vrouwen en statistisch significant bleef. Voor mannen was het verband tussen de consumptie van bewerkt vlees en PTO-risico niet langer statistisch significant, hoewel het verband nog steeds positief was. Er werden geen associaties gevonden tussen rood vlees (algemeen), gevogelte of visconsumptie en PTO-risico.

De familiegeschiedenis van kanker bleek in ons onderzoek geen onafhankelijke risicofactor te zijn (zie hoofdstuk 6). We zagen wel een matig verhoogd PTO-risico bij deelnemers die een broer of zus met kanker rapporteerden in vergelijking

met degenen die dat niet deden, en we vonden een licht verhoogd PTO-risico bij deelnemers met een familiegeschiedenis van kanker bij een zus. Er werd geen associatie gezien voor ouders of deelnemers met een broer met een familiegeschiedenis van kanker. PTO bleek niet geassocieerd te zijn met een familiegeschiedenis van borst-, eierstok-, endometrium-, darm-, maag-, long, prostaat-, blaas-, pancreas-, hoofd-halskanker, lymfoom en/of leukemie. We hebben wel een verminderd PTO-risico waargenomen bij deelnemers die een familiegeschiedenis van nierkanker meldden, hoewel dit slechts op drie gevallen was gebaseerd. Deze bevinding kan daarom een toevalsbevinding zijn.

In onze studie werd een niet-significante positieve associatie tussen T2DM-status en PTO-risico waargenomen (zie hoofdstuk 7). In de naar geslacht gestratificeerde analyse bleek dat de associatie sterker werd bij vrouwen. Deelnemers die bij de diagnose van T2DM <50 jaar oud waren, bleken een statistisch significant verhoogd PTO-risico te hebben, dat opnieuw duidelijk sterker werd bij vrouwen.

In een andere studie hebben we onderzocht of het naleven van de WCRF/AICR leefstijlaanbevelingen voor kankerpreventie geassocieerd was met PTO-risico. We onderzochten naleving van het advies met betrekking tot lichaamsvet, lichamelijke activiteit, plantaardige voeding, vleesconsumptie en alcohol. We zagen dat deelnemers met de hoogste naleving van de aanbevelingen een statistisch significant verminderd PTO-risico hadden in de leeftijd- en geslachtsgecorrigeerde analyse (zie hoofstuk 8). In de multivariabele analyse zagen we dat de associatie tussen het opvolgen van de aanbevelingen en het PTO-risico niet langer statistisch significant was na aanvullende correcties voor rookgedrag. Deelnemers met de hoogste naleving van de aanbevelingen voor vlees (rood en bewerkt vlees) en alcoholconsumptie bleken een statistisch significant verminderd PTO-risico te hebben. Naleving van de aanbevelingen met betrekking tot lichaamsvet, lichamelijke activiteit of inname van plantaardige voeding was niet geassocieerd met PTO-risico.

In ons uitgebreide review hebben we systematisch gezocht naar epidemiologische studies over mogelijke PTO-risicofactoren (zie hoofdstuk 9). Het bestaande epidemiologische bewijs beschrijft associaties tussen roken, familiegeschiedenis van kanker, diabetes mellitus, middelomtrek, en immuniteitsstoornissen in relatie tot PTO-risico, terwijl er zwakkere associaties werden gevonden voor alcoholconsumptie, opleidingsniveau, en geen associaties werden gevonden

voor de inname van dierlijke of plantaardige voeding. Om de risicofactoren die zijn waargenomen in de NLCS-onderzoeken en het bestaande epidemiologische bewijs te evalueren, hebben we de beoordelingscriteria gebruikt als bewijs voor kankerpreventie zoals gerapporteerd door het WCRF; variëren van overtuigend tot beperkt-geen conclusie-bewijs. Door deze beoordelingscriteria toe te passen, lijkt roken een gevestigde verhoogde risicofactor voor PTO te zijn, terwijl er beperkt suggestief bewijs werd gevonden voor alcoholconsumptie, diabetes mellitus, en familiegeschiedenis van kanker, terwijl er geen overtuigende associaties werden gevonden voor antropometrie, inname van voeding (dierlijk of plantaardig), immuniteitsstoornissen, levensstijl (algemeen), lichamelijke activiteit of sociaaleconomische status in relatie tot PTO-risico.

Dit proefschrift wordt afgesloten met een samenvatting van de belangrijkste bevindingen, methodologische overwegingen en toekomstige aanbevelingen, en conclusie. Over het algemeen lieten de onderzoeken zien dat roken, alcoholconsumptie, diabetes mellitus, en familiegeschiedenis van kanker de belangrijkste associaties zijn met betrekking tot het ontwikkelen van PTO. Concluderend blijkt het aanhouden van een gezonde leefstijl gunstig te zijn bij het voorkomen van PTO, wat van groot belang is omdat de ziekte gepaard gaat met een sombere prognose.

Impact

CUP is a complex malignancy due to its heterogeneous nature, classification and vulnerable patient group. The prognosis for CUP patients is bleak and the great majority of patients do not survive one year after diagnosis. In general, cancer treatment(s) is targeted on the primary tumour origin, however, due to the inability of identifying the origin, it may be more beneficial to focus on disease prevention. Globally, approximately 42-50% of cancers could be prevented if modifiable risk factors are being addressed appropriately (1). Hence, the identification of risk factors associated with CUP may guide preventative methods. To acquire knowledge on risk factors that are associated with CUP, we investigated risk factors that have been associated with cancers of known primary sites such as alcohol consumption, cigarette smoking, anthropometry, physical activity, vegetable and fruit consumption, meat consumption, family history of cancer, and diabetes mellitus. These risk factors are investigated as individual components but also as an overall lifestyle. The findings of our individual component investigations indicate that cigarette smoking, alcohol consumption (dose-response relationship), and the consumption of beef and processed meat (predominantly in women) are associated with increased CUP risk in the NLCS. Positive associations were also seen for family history of cancer in siblings, and diabetes mellitus, although these were statistically non-significant. No associations were observed between anthropometry, physical activity, or vegetable and fruit consumption in relation to CUP risk. Findings of our overall lifestyle investigation indicate that adhering to lifestyle recommendations for cancer prevention applies to reducing CUP risk as well. To compare our findings with the existing epidemiological evidence on CUP risk factors we have written a comprehensive review that could be included in future evidence-based guidelines for cancer prevention of CUP. Our study findings may be useful for the Continuous Update Project which is continuously updated by the World Cancer Research Fund (WCRF) and American Institute for Cancer Research (AICR), in which cancer prevention recommendations, as well as public health and policy implications concerning lifestyle factors in relation to cancer risk, are systematically reviewed (2). Scientific evidence concerning CUP is not yet included in their summary of evidence as there was insufficient data available, consequently, our results together with the existing epidemiological evidence as discussed in our comprehensive review may be included in future Continuous Update Projects. In addition, to raise more awareness about CUP and its associated risk factors, knowledge sharing is essential.

In 2012, the first oncology guidelines for CUP patients were published in the Netherlands, which became the start of all developments for CUP patients today. Since 2014, various patient organisations put their effort into raising CUP awareness. In the course of this PhD-project, patient organisations from the United Kingdom (CUP Foundation Jo Symons), Ireland (Sarah Jennifer Knott Foundation) and the Netherlands (Missie Tumor Onbekend) organised the first World CUP Awareness Week in September 2021. We were very pleased to have received the opportunity to disseminate our study findings at this conference. In the Netherlands specifically, Missie Tumor Onbekend and the Dutch Federation for Cancer Patients Organisation (Nederlandse Federatie van Kankerpatiënten Organisaties) put effort into collaborations with the Dutch Society for Medical Oncology (Nederlandse Vereniging voor Medische Oncologie), Dutch Society for Pathology (Nederlandse Vereniging voor Pathologie), Hartwig Medical Foundation and Dutch health insurers. Those collaborations aim to standardise the necessary diagnostic examinations and to shorten the diagnostic process into a maximum length of 6 weeks, alongside the accessibility to advanced molecular diagnostics for all patients regardless of the hospital where the patient was initially examined. In October 2021, a national multimedia campaign was released in the Netherlands which promoted CUP awareness on a considerable scale. This increased awareness was especially valued by the patient organisation, as it gave recognition to the disease, but also indicated the need for action. Another important asset in this PhD-project, are the collaborations with external researchers: Caroline Loef from the Netherlands Cancer Registry, Fatemeh Kazemzadeh and Iris Nagtegaal from the Department of Pathology at the Radboud University Medical Centre, and contributors: Warnyta Minnaard and Francine van der Heijden from the patient organisation Missie Tumor Onbekend. Together with their inputs, we were able to write a comprehensive review of CUP risk factors. Due to the increased CUP awareness, and networking between researchers and medical doctors, there has been a start of outpatient clinics for CUP patients throughout the country. The first outpatient clinics were opened in the Netherlands Cancer Institute/Antoni van Leeuwenhoek hospital (Amsterdam) and Erasmus University Medical Center (Rotterdam), and more hospitals are following. Until now, the utilisation of whole-genome sequencing (WGS) is not part of routine care yet, due to improper arrangements for reimbursements. This technique is particularly useful to guide tumour-targeted treatments and therefore CUP patients may substantially benefit from its perspectives. The Dutch Federation for Cancer Patients Organisation is

advocating for more rapid implementations of DNA techniques as well as proper reimbursements of these investigations, and in which hospitals the investigations are being conducted. The initiation of the outpatient clinics enlarges the opportunity for CUP patients to be directed to the best possible care. The wider availability of WGS and proper reimbursements would make it easier for CUP patients to receive treatment perspectives and therapies in nearby located hospitals.

To further disseminate CUP knowledge on an international scale, the International Agency for Research on Cancer (IARC), as part of the World Health Organisation (WHO) of the United Nations (UN) could play an important role in further extending knowledge on CUP by disseminating the findings of our investigations, as their organisation influences global cancer control policies (3). On a national scale, it is beneficial to share our study findings with informative news outlets such as Koningin Wilhelmina Fonds (KWF)-kankerbestrijding, Wereld Kanker Onderzoek Fonds, Volksgezondheid en Zorg, and voeding&kankerinfo.

Knowledge transfer

The scientific evidence obtained from this study has been published in peer-reviewed journals. All articles were published open access to maximise accessibility to its results at no cost. In addition, results were presented at international and national scientific conferences such as the GROW-Science Day in Maastricht, the Netherlands (2019, 2020, 2021), in the masterclass Nutrition and Cancer in Wageningen, the Netherlands (2019), at the virtual conference on Cancer Prevention organised by the German Cancer Research Center (2020), at the Dutch Epidemiological Conference (WEON) (2020, 2021), at the virtual annual meeting of the American Association for Cancer Research (2021), at the virtual World CUP Awareness Week (2021), and Science Day MUMC+ (2021). Findings and interviews have also been shared on several national news outlets such as the Netherlands Cancer Registry, the Dutch World Cancer Research Fund (Wereld Kanker Onderzoek Fonds), and patient organisations (Missie Tumor Onbekend & Patiëntenplatform Zeldzame Kankers); to raise more awareness about the disease and its associated risk factors.

Conclusion

In this thesis, we have investigated various risk factors in relation to CUP risk. The results of our studies can be used by other researchers as it contributes to the current epidemiological evidence of CUP. Within the time frame of this PhD project, a lot of CUP awareness has been raised both internationally and nationally. This increased awareness has brought together essential stakeholders that can have a major influence on the prognosis for future CUP patients. We hope that, together with these stakeholders, our work can contribute to reducing the disease occurrence.

References

1. Wild CP, Bucher JR, De Jong BW, Dillner J, Von Gertten C, Groopman JD, et al. Translational cancer research: balancing prevention and treatment to combat cancer globally. J Natl Cancer Inst. 2015;107(1):1-5.
2. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. 2018.
3. International Agency for Research on Cancer. IARC's Mission: Cancer research for cancer prevention [Available from: https://www.iarc.who.int/about-iarc-mission/.

Dankwoord

Gedurende mijn promotietraject heb ik mij kunnen ontwikkelen als onderzoeker, maar ook als persoon. Ik ben dan ook erg dankbaar voor de mogelijkheden die hiervoor beschikbaar waren en voor de steun van de mensen om mij heen.

Promotieteam

Allereerst wil ik de leden van mijn promotieteam, *Leo, Piet en Caroline*, bedanken. Zonder jullie bijdragen en feedback zou het niet gelukt zijn om tot dit mooie eindresultaat te komen.

Leo, bedankt voor je vertrouwen om mij aan te stellen als promovenda voor dit onderzoek. De afgelopen vier jaar zijn voorbij gevlogen. Ik heb met heel veel plezier, samen met jou, aan dit project gewerkt. We hebben ontzettend veel digitale meetings gehad, desondanks heb ik het als een hele prettige samenwerking ervaren. Je stond altijd voor mij klaar als dat nodig was, of dit nou werkgerelateerd of om privésituaties ging, daar ben en blijf ik je enorm dankbaar voor. We hebben veel leuke en diepgaande discussies gehad over het onderzoek, maar ook over egels en vleermuizen in de tuin. In de meeste overleggen hadden we genoeg te bespreken, maar het kwam ook wel eens voor dat er weinig agendapunten op de planning stonden. We dachten dan dat we met deze overleggen het snelst klaar zouden zijn, maar de tijd heeft geleerd dat we in die gevallen vaak zelfs uitlooptijd wisten te behalen. Mede doordat we allebei door enthousiasme bleven praten, vooral als het ging om Leudal en de Meinweg. Ik kijk ook terug op een gezellige avond toen ik met Jeroen mocht langskomen bij jou thuis, waar we je vrouw mochten ontmoeten en waar jij je heerlijke kook-skills hebt gepresenteerd! Ook zijn we met de trein naar de CUP-conferentie in Londen gegaan waar we, samen met Caroline en Laura van het IKNL, kennis hebben gemaakt met een aantal internationale collega's. Graag wil ik je bedanken voor de mogelijkheden die je mij hebt geboden gedurende het promotietraject, waarbij ik een aantal cursussen mocht volgen, conferenties mocht bijwonen en waarin ik met veel plezier stagiair(e)s mocht begeleiden. Je motivatie en steun daarin waren onmisbaar!

Piet, ik wil je graag bedanken voor jouw steun, evenals je kritische en waardevolle feedback. Jouw deskundigheid heeft ertoe geleid dat we de kwaliteit van de onderzoeken tot een hoger niveau konden brengen. Ik heb dan ook veel van je

geleerd als onderzoeker en neem dit in mijn carrière mee. Ook wil ik je bedanken voor de ervaring om tutor te mogen zijn in het masterblok Observational Research, zowel als tutor en student zijnde heb ik daar heel veel van geleerd. Ik deel nog steeds de mening dat de NLCS een hele mooie studie is en ik ben ook nog steeds heel dankbaar dat ik deel heb mogen uitmaken van het team.

Caroline, ook jou wil ik graag bedanken voor onze prettige samenwerking. Ik heb ontzettend veel van je geleerd, middels je feedback die vanuit een ander oogpunt kwam (kankerregistratie), maar ook van je enthousiasme om PTO op de kaart te zetten en niet geheel onbelangrijk; van jouw hartelijkheid als mens! Ik bewonder het enorm hoe jij in de afgelopen jaren de handen uit de mouwen hebt gestoken om samen met de patiëntenvereniging van PTO, aandacht te vragen voor de ziekte zowel nationaal als internationaal. Dankzij jouw inzet om de juiste 'koppen bij elkaar te brengen' zoals je dit vaak zo mooi benoemd, weet ik zeker dat er nog veel resultaten te behalen zijn voor PTO-patiënten. Ook al werk jij vanuit het noorden van het land (Friesland), was jij met je gedachte vaak ook in het zonnige zuiden (Limburg), of het nou ging om natuurbrand, overstroming, of steun in moeilijke tijden, jij was er altijd! Enorm bedankt!

Beoordelingscommissie

De leden van de beoordelingscommissie, prof. V. Tjan-Heijnen, prof. B. Kremer, prof. J. Muris, prof. V. Lemmens, en prof. V. Smit, wil ik bedanken voor het kritisch lezen en het beoordelen van dit proefschrift.

A

Co-auteurs/samenwerkingen

Rob, ik wil je graag bedanken voor de prettige samenwerking en voor het ontvangen van jouw feedback. Ondanks dat je gedurende het promotietraject met pensioen bent gegaan, heb je toch nog meegewerkt aan de studies, daar heb ik veel waardering voor! Hopelijk kan je na de afronding van dit project nóg meer genieten van je pensioen.

Fatemeh and Iris, I am very pleased that we got the opportunity to collaborate for our review. It has been a great pleasure to learn from you. I am very convinced that your research into artificial intelligence and pathology will improve the current CUP knowledge. I am already looking forward to read your future articles!

I would like to give my thanks to *Gabriel,* you have helped our team a lot by checking the correct use of the English language, but also by making critical comments on our manuscripts. Both improved the quality of our papers. Besides work, you were always there to talk if needed, which was very much appreciated!

Interns

I would also like to give my thanks to the interns that have contributed their time and effort to our CUP research. *Alexander*, it was very interesting to study the relationship between a family history of cancer and CUP. Although most of our meetings were on Skype, it was very pleasant to work with you. *Sharmi*, I am very thankful for our collaboration to study diabetes mellitus in relation to CUP risk. I have learned a lot about your enthusiasm and hope to pursue it in my further career. *Anna*, I very much enjoyed working together with you on studying the relationship between socioeconomic status and CUP. It was very nice to conduct this work while having a mutual interest in global health.

Collega's

De collega's van de afdeling Epidemiologie wil ik graag bedanken voor de leuke tijd op de afdeling. Er is genoeg gebeurd in de afgelopen jaren om nog lang naar terug te kijken (voedselvergiftiging, cyber hack en vooral de COVID-19 uitbraak). Desondanks was het gezellig om met jullie (o.a.) te borrelen en pogingen te doen om uit de digitale escaperoom te ontsnappen (helaas is dit beide keren niet gelukt). De *dames van het secretariaat* wil ik bedanken voor hun hulp en administratieve ondersteuning. *Jolanda,* regelmatig had ik je nodig om afspraken in te plannen, geen vraag was te moeilijk en je nam hiervoor altijd uitgebreid de tijd, dankjewel! *Conny*, ook jou wil ik graag bedanken, als er iets geregeld moest worden stond jij altijd met open armen voor mij klaar. Ook was het gezellig om je zo nu en dan tegen te komen in Heugem of in Randwyck! *Jos,* of het nou ging om een muismat, toetsenbord, of over reistips; geen vraag was te gek en je nam altijd de tijd voor mij, bedankt hiervoor! Het was altijd gezellig om met je te kletsen! Ook als er ICT-problemen waren, wisten *jij, Harry en Ron*, altijd iets te regelen. Zelfs op de korte termijn of op afstand.

Lieve *Christel*, ik wil jou graag bedanken voor je luisterend oor, je motiverende stem en onze gezellige wandelingen. Iedere keer wist je me te verrassen met een nieuwe

mooie route, ik hoop dan ook dat we onze wandelingen nog voort kunnen zetten in de toekomst! Bij zowel ups en downs stond jij voor me klaar, dit was al zo tijdens de master en heeft zich verder ontwikkeld tijdens de PhD. Je bent een schat!

Colinda, Adri en Tanja wil ik graag bedanken voor de gezellige gesprekken op de afdeling, als we elkaar tegenkwamen namen jullie altijd de tijd om even bij te kletsen.

Verder wil ik de (oud)-aio's, bedanken voor de gezelligheid tijdens de (kerst) lunches, wandelingen, borrels, en etentjes. Mede dankzij deze ervaringen kijk ik terug op een leuke PhD-tijd, ondanks dat we elkaar zelfs 2 jaar moesten missen door de COVID-19 uitbraak.

Jacqueline, bedankt voor onze gezellige gesprekken! Jij was altijd bereid om een leuk praatje aan te gaan of het nou in de wandelgang, de sportschool of het keukentje van Deb was! Ik heb veel van jou geleerd en hoop dat we nog veel mooie feestjes kunnen vieren!

Sophie, ook jou wil ik graag bedanken voor onze fijne gesprekken, gezellige wandelingen, energievolle spinning-lessen, en mountainbike avonturen die we hopelijk nog lang kunnen voortzetten zowel in Maastricht, als in de omgeving van Amsterdam!

Annaleen, bedankt voor je luisterend oor en de gezellige gesprekken tijdens onze wandelingen, maar ook tijdens de leuke Indiase etentjes en avonturen in Gent! Ook als je in de omgeving Maastricht bent, denk je altijd even aan mij, wat ik heel tof vind!

Maya, ontzettend bedankt voor onze gezellige etentjes en gesprekken, maar ook voor onze Skype-sessies die we na de afronding van jouw PhD hebben doorgezet! Ik vind het super leuk dat we elkaar nog af en toe zien in Eindhoven of Maastricht.

Marlou Floor, Josien en Kelly, bedankt voor de gezellige momenten die wij samen hebben beleefd gedurende deze 4 jaar. Ik zal onze gesprekken, etentjes, kerstmarkt en pretpark-avonturen zeker niet vergeten en ben jullie daar dankbaar voor! *Marlou Floor*, bedankt voor onze gezellige choco-momentjes en de gezellige tijd in onze kantoorruimte. *Josien*, zelfs bedankt voor het delen van die ene bitterbal, ik zal je eraan blijven herinneren! *Kelly,* het zien van jouw fotoalbums met betrekking tot reizen was een motiverende inspiratie om de volgende reis uit te stippelen!

Lloyd, Maya, Lisette en Jeroen, ik wil jullie bedanken voor het introduceren van de wandelroute rondom Deb. Ook wil ik jullie bedanken voor jullie bereidheid om te helpen als STATA-hulplijn bij het uitvoeren van de NLCS-analyses. *Jeroen,* jouw vrolijkheid en enthousiasme waren absoluut onmisbaar. In het kader van 'heb je even voor mij' was jij er altijd!

Ariane, Jessica en Linda, het was altijd gezellig om met jullie te kletsen. Het lukte nooit om een kort gesprek te houden omdat we te veel wilden bespreken. Ik kijk er met plezier op terug!

Vrienden

Daarnaast wil ik ook mijn toffe vrienden bedanken, zonder jullie steun en toeverlaat zou het mij niet gelukt zijn om dit promotietraject succesvol af te ronden. Jullie waren er zowel in de positieve als de negatieve momenten, ook al waren er helaas erg veel negatieve momenten. Mede dankzij jullie geniet ik nog steeds van de tijd in Mestreech!

Juliet (Juul), helaas woon je erg ver weg, toch ben je ook altijd dichtbij. Je bent een geweldige vriendin die mij altijd weet op te fleuren, je bent absoluut onmisbaar! Bedankt voor jouw altijd aanwezige steun en toeverlaat! Ik hoop dat iedereen een 'Juul' in zijn of haar leven mag ontmoeten. Ik kijk uit naar onze toekomstige avonturen, samen met jullie meisjes! Dider, enorm bedankt voor de ABN-checks!

Marly en *Nina*, lieve dushis, ook wij wonen helaas niet meer dicht bij elkaar, toch voelt het altijd als vanouds als we weer samen zijn, bedankt voor jullie steun en gezellige gesprekken, telefoontjes en interesse die jullie hebben getoond in het onderzoek! Jullie stonden altijd voor mij klaar! Dankjewel!

Stevie, bedankt voor de vertraging... en voor de gezellige gesprekken, sportsessies!!!, kook-, puzzel-, en spelavonden. Het was heel fijn om de PhD-ervaringen met jou te kunnen uitwisselen en om de frustraties los te laten tijdens de bodyattack lessen!

Joyce, wijfie, bedankt voor de gezellige momenten die ik samen met jou in de afgelopen jaren heb mogen beleven, of het nou in Zaandijk of in Maastricht was! Ook hebben we hele leuke vakanties mogen meemaken. We kennen elkaar nog niet eens zo lang, maar het voelt alsof ik je al jaren ken. Je weet er altijd voor me te zijn, ik waardeer het enorm.

Danny, we kennen elkaar al van kleins af aan, waardoor je eigenlijk een tweede broer bent. Ook jij bent er altijd voor mij geweest in zowel de leuke als minder leuke momenten, zowel in Limboland als op vakanties!

Bea, a few months ago you became my new flatmate, but also my new friend. You have been amazing to discuss work and social life with, your happiness and enthusiasm are amazing. I am very happy that we got to meet, and I am very much looking forward to all the activities that you will plan for us to do!

De volgende dames van het sporten (met ongekende energie) wil ik bedanken voor de fijne en nodige afleiding na werktijd. *Erika,* ik ben jou enorm dankbaar voor je luisterend oor en je positieve kijk op dingen! Dankzij jou haal ik extra veel plezier uit de bodyattack lessen en heb ik me kunnen ontwikkelen als instructeur. Ook wil ik *Sarina en Caroline* bedanken*,* jullie motivatie en enthousiasme om te sporten, evenals onze fijne gesprekken zijn onmisbaar! Ik hoop dat we nog veel leuke en sportieve momenten gaan beleven!

Familie

Graag wil ik mijn ouders *Edy en Lily*, en mijn broer *Daan* bedanken. Het was niet altijd een makkelijke tijd, toch kon ik altijd bij jullie terecht als dit nodig was zowel in Herkenbosch als in Maastricht. Jullie hebben mij altijd gesteund bij het doorzetten van mijn studies en carrière. Nu kunnen jullie trots zijn op jullie 'moes' die uiteindelijk doctor is geworden!

Als laatste wil ik iemand bedanken die voor mij heel bijzonder is. Lieve *Umit*, we hebben elkaar een aantal maanden geleden leren kennen. De afgelopen maanden zijn behoorlijk turbulent geweest, maar ook ontzettend fijn. Je bent een bijzonder persoon, ik hoop dat we er 'saampjes' een mooie toekomst van gaan maken.

About the author

Karlijn Hermans was born on the 6th of May 1994, in Roermond, the Netherlands. After graduating from secondary school, Karlijn studied Medical Imaging and Radiation Therapy at Fontys University of Applied Sciences in Eindhoven, with specialisations in Radiology and Radiotherapy (2011-2015). Within this bachelor program, she completed a minor in Management in Healthcare at the HAN University of Applied Sciences in Nijmegen (2013-2014). For her bachelor thesis, she investigated volume and density changes in the liver, spleen, and kidneys up to 36 hours post-mortem at the Radboud UMC Nijmegen (2015).

Karlijn pursued her career in research by following a premaster in Biomedical Sciences at the Radboud University in Nijmegen (2015-2016). After, she started the master program in Global Health at Maastricht University, for which she investigated determinants of perinatal mortality in Suriname (2016-2017). To acquire additional knowledge of research methodology and statistics, she continued her studies in the master program of Epidemiology at Maastricht University for which she examined the diagnostic accuracy and radiation dose of 4D-Dual Energy Computed Tomography for the localisation of parathyroid adenoma (2017-2018).

After the completion of her masters, Karlijn started her PhD at the Department of Epidemiology of Maastricht University (GROW-School for Oncology and Developmental Biology) (2018-2022). She investigated risk factors of Cancer of Unknown Primary by using data from the Netherlands Cohort Study on diet and cancer. Scientific results of this PhD-project, which are presented in this thesis, have been published in international peer-reviewed journals and were presented at national and international conferences.

List of publications

Hermans, KE, van den Brandt, PA, Loef, C, Jansen, RL, Schouten, LJ (2021). Alcohol consumption, cigarette smoking and cancer of unknown primary (CUP) risk: Results from the Netherlands cohort study. *Int. J. Cancer,* 1-12.

Hermans, KE, van den Brandt, PA, Loef, C, Jansen, RL, Schouten, LJ (2020). Anthropometry, physical activity and cancer of unknown primary (CUP) risk: Results from the Netherlands cohort study. *Cancer Epidemiol.* 69.

Hermans, KEPE, van den Brandt, PA, Loef, C, Jansen, RLH, Schouten, LJ (2022). Vegetable and fruit consumption and cancer of unknown primary (CUP) risk: Results from the Netherlands cohort study. *BMC Cancer.* 22: 399.

Hermans, KEPE, van den Brandt, PA, Loef, C, Jansen, RLH, Schouten, LJ (2021). Meat consumption and cancer of unknown primary (CUP) risk: Results from the Netherlands cohort study. *Eur. J. Nutr.*

Grewcock, AL, **Hermans, KE**, van den Brandt, PA, Weijenberg, MP, Loef, C, Jansen, RL, Schouten, LJ (2021). Family History of Cancer in First Degree Relatives and Risk of Cancer of Unknown Primary. *Eur. J. Cancer Care,* 1-9.

Hermans, KE, Haque S, van den Brandt, PA, Loef, C, Jansen, RL, Schouten, LJ (2022). Type 2 diabetes mellitus and cancer of unknown primary (CUP) risk: results from the Netherlands cohort study. *Eur. J. Cancer Prev.*

Hermans, KEPE, van den Brandt, PA, Loef, C, Jansen, RLH, Schouten, LJ (2022). Adherence to the World Cancer Research Fund and the American Institute for Cancer Research lifestyle recommendations for cancer prevention and cancer of unknown primary risk. *Clin. Nutr.* 41: 526-535.

Hermans, KEPE, Kazemzadeh, F, Loef, C, Jansen, RLH, Nagtegaal, ID, Van den Brandt, PA, Schouten, LJ (2022). Risk factors for Cancer of Unknown Primary: a literature review. *Submitted for publication.*